PANTHEON

Community-Based Smart City Digital Twin Platform

for Optimised DRM operations and Enhanced Community

Disaster Resilience

# D3.5

## ETHICAL, LEGAL, AND SOCIETAL PANTHEON DESIGN CONSIDERATIONS

The project has received funding from the European Union’s Horizon Europe programme under Grant Agreement N°101074008.

# DOCUMENT INFO

| | |
|---|---|
| ***Deliverable Number*** | D3.5 |
| ***Work Package Number and Title*** | WP3 – Ethical, legal, and societal PANTHEON design considerations |
| ***Lead Beneficiary*** | ISPC |
| ***Due date of deliverable*** | 31/12/2023 (M12) |
| ***Deliverable type[1]*** | Document, report |
| ***Dissemination level[2]*** | PU - Public |
| ***Author(s)*** | Kiril Shtefchyk (ISPC), Lola Vallès (ISPC), Iacob Crucianu (SIMAVI), Otilia Bularca (SIMAVI), Nick Petropoulos (HPOL) |
| ***Internal reviewer(s)*** | Mike Karamousadakis (THL), Constanze Geyer (JOAFG), Otilia Bularca (SIMAVI), Ana-Maria Dumitrescu (SIMAVI) |
| ***Version - Status*** | V1.4 |

# TASK ABSTRACT

The present deliverable "Ethical, legal and societal design considerations" is the result of Task 3.5. It aims to identify and analyse potential ethical, social and legal issues that may arise during the design, development and deployment of the PANTHEON system. Based on this study, a number of recommendations have been extracted that architects, designers, developers and end-users should consider throughout the system's lifecycle.

[1] **Please indicate the type of the deliverable using one of the following codes**:
R = Document, report
DEM = Demonstrator, pilot, prototype, plan designs
DEC = Websites, patents filing, press & media actions, videos
DATA = data sets, microdata
DMP = Data Management Plan
ETHICS: Deliverables related to ethics issues.
OTHER: Software, technical diagram, algorithms, models, etc.

[2] **Please indicate the dissemination level using one of the following codes**:
PU = Public
SEN = Sensitive

## REVIEW HISTORY

| | | | |
|---|---|---|---|
| **Version** | **Date** | **Modifications** | **Editor(s)** |
| **1.0** | 22.12.2023 | First draft | Kiril Shtefchyk & Lola Vallès (ISPC) |
| **1.1** | 03.01.2024 | Internal Review | Mike Karamousadakis (THL) |
| **1.2** | 09.01.2024 | Internal Review | Constanze Geyer (JOAFG) |
| **1.3** | 16.01.2024 | Internal Review | Otilia Bularca & Ana-Maria Dumitrescu (SIMAVI) |
| **1.4** | 18.01.2024 | Final version | Kiril Shtefchyk & Lola Vallès (ISPC) |

## DISCLAIMER

The document is proprietary of the PANTHEON consortium members. No copying or distributing, in any form or by any means, is allowed without the prior written agreement of the owner of the property rights.

Funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# EXECUTIVE SUMMARY

Within WP3 that sets the requirements for the design process of PANTHEON system, D3.5 deals with ethical, legal and societal issues related to this design process and the future PANTHEON system. Therefore, this report is the result of Task 3.5 "PANTHEON Ethical, Legal and Societal Considerations", which aims to provide a set of ethical and legal recommendations to be considered in the design of the PANTHEON system.

Chapter 2 presents the legal framework, including certain human rights and the main legal texts in which they are embedded. Within human rights, the challenges that new technologies have posed are described, including AI systems. Emergency situations are discussed as special contexts where rights are not absolute and in which they can be restricted by public authorities, if there is a greater public good at stake and this is done responsibly and proportionately.

Chapter 3 presents PANTHEON's ethical approach to the technologies that will serve as the basis for the creation of a planning and training tool to support relevant disaster risk management actors. On the one hand, the main technologies to be used by PANTHEON are highlighted. The intended use of SCDTs, IoT systems, AI, DSS and risk monitoring systems is briefly explained. On the other hand, the second section addresses the ethical, legal and social issues that may arise in the development of these technologies. The various issues are presented in a general way, highlighting risks, threats, and some ways of dealing with them. Emphasis is placed on security and privacy, but also on other issues such as equality, non-discrimination, justice, freedom of expression or assembly.

Chapter 4 deals with PANTHEON's security and privacy by design framework, covering data security, privacy, ethics, social acceptability and legal aspects. Security criteria are formulated in terms of confidentiality, integrity and availability, the CIA triad.

Finally, as a main result of this Deliverable, a set of concrete recommendations on ethical, legal and social issues that PANTHEON’s system design process will follow are gathered in a catalogue and presented in chapter 5.

# 1 INTRODUCTION

We live in a time of rapid change, especially when it comes to technology. In just a few decades, technologies have been developed that are revolutionising the way we live. We cannot deny that technological advances have made many things in our lives easier. Inventions like the Internet allow us to communicate with other people at any time. Drones can access disaster sites and provide valuable information without putting people's lives at risk. Virtual reality and simulations allow for complex disaster training that would be economically and logistically impossible in real life. Recently, Artificial Intelligence (AI) has emerged as the quintessential revolutionary technology. However, technology is not an independent and isolated element of society. We cannot ignore the fact that any technological innovation will have legal and social implications that need to be adequately addressed. If we allow technology to advance without paying attention to its impact on issues such as security, privacy, justice, responsibility or equality, the price of the risks will be too high for the benefits, which are not few.

The overall objective of PANTHEON is precisely to exploit technological innovations such as the Smart City Digital Twin (SCDT), the Internet of Things (IoT) and AI to provide a useful planning and training tool for relevant actors in disaster situations. The purpose of such a tool is to enable informed decision-making by supporting different stakeholders. Also, the PANTHEON system will have ethical, legal, and social implications. The purpose of this report is therefore to analyse these impacts and to propose recommendations for the architects and designers of the platform to create an ethical system. It should be noted that at the time of writing, the specific elements, features, and details of PANTHEON have not yet been finalised. While it is true that preliminary ideas and elements have been proposed, the project and the system may change in the near future. Consequently, this report is preliminary in nature and is intended to guide those responsible for the design of the system in the creation of a tool that is useful, yet compliant with current legislation and ethical codes. The result is a catalogue of recommendations, which the reader will find in Section 5 of the document.

However, the effort to integrate ethics into the system does not end with this deliverable. The partners responsible for the design, development and deployment of the PANTHEON system will work on the aspects mentioned here in their respective tasks, and the ethical and legal issues of the system will be addressed in future deliverables as the elements and features of the platform are better defined and the risks and threats can be more clearly identified. The present report will serve as a basis and guide for them to address in more depth and detail the issues mentioned here, as well as others identified along the way and not foreseen in this document. Finally, it should be emphasised that no one can foresee all the consequences of a technological innovation. However, this should not be an obstacle to thinking about and addressing risks from the design phase of a technological system, thus delivering a product based on ethics by design, including security and privacy.

# 2. LEGAL FRAMEWORK

## 2.1. HUMAN RIGHTS

The European Union (EU), and by extension all its Member States, is deeply committed to human rights. This is enshrined in Art. 2 of the Treaty on European Union (EU, 2016b), which states that "the Union is founded on the values of respect for human dignity, freedom, democracy, equality, the rule of law and respect for human rights, including the rights of persons belonging to minorities". As mentioned in several deliverables, particularly those related to the ethics of this project, the PANTHEON consortium partners are also committed to respecting fundamental rights in all their activities. Therefore, the outcome of the PANTHEON project - i.e., a digital system/platform to promote a sense of community among the different groups of citizens, volunteers and organisations involved in disaster management, and to facilitate effective resource mobilisation - is designed to respect the rights of individuals and communities, and thus represents a community-based and fundamental rights-respecting tool.

Several international and European texts recognise and enshrine the rights of the individual, the cornerstone of which is the Universal Declaration of Human Rights (United Nations, 1948). Article 1 of the Declaration states that "all human beings are born free and equal in dignity and rights...", thus recognising the freedom and equality of the individual. Closely linked to this right is the prohibition of discrimination, set out in Article 7, which states that "All are equal under the law and are entitled without any discrimination to the equal protection of the law. All shall be entitled to equal protection against any discrimination in violation of this Declaration and against any incitement to such discrimination." Art. 26 of the International Covenant on Civil and Political Rights (UN, 1966a) also refers to equality, stipulating that "All persons are equal under the law and are entitled without any discrimination to the equal protection of the law". This protection is provided irrespective of "race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status". In this regard, it is also relevant to mention Art. 15.1 of the International Covenant on Economic, Social and Cultural Rights (UN, 1966b), according to which all States Parties recognise the right of everyone to "enjoy the benefits of scientific progress and its applications".

Similarly, Article 14 of the European Convention on Human Rights (Council of Europe, 1950) recognises the prohibition of discrimination in the enjoyment of rights and freedoms. Article 8 of the Treaty in the Functioning of the EU (European Union, 2016c), one of the Union's fundamental legal texts, refers to the fight against all forms of discrimination by all its Member States and to the elimination of inequalities between men and women. Similarly, the Charter of Fundamental Rights of the European Union (EU, 2016a) aims to guarantee equality between men and women (Art. 23), while Art. 21 prohibits any form of discrimination. It places particular emphasis on the rights of the child (Article 24), the rights of the elderly (Article 25) and the integration of people with disabilities (Article 37).

In this context, it is important to distinguish between individual rights and the collective rights of certain particularly vulnerable groups, the latter being a development of the former. One of the relevant instruments in this regard is the Convention on the Elimination of All Forms of Discrimination against Women (UN, 1979), which establishes the basic principles of gender equality and non-discrimination. States parties to the Convention must commit to taking concrete measures to ensure gender equality in their legislation, policies, and practices. Another specially protected group are children, whose rights are recognised in the Convention on the Rights of the Child (UN, 1989), which is based on four fundamental principles: (a) non-discrimination, (b) the best interests of the child, (c) the right to life, survival, and development, and (d) their active participation in matters affecting them. The Convention establishes a wide range of rights for children (i.e., the right to life, to a name and nationality, not to be separated from their parents unless it is in their best interests, to education, health, protection from abuse and exploitation, participation in decisions affecting them) and requires states parties to take measures to ensure respect for and protection of these rights, with the Committee on the Rights of the Child responsible for monitoring their implementation in the various States Parties. Another relevant text is the Convention on the Rights of Persons with Disabilities (UN, 2006), which aims to promote, protect, and ensure the full and equal enjoyment of all human rights and fundamental freedoms of persons with disabilities. Its basic principles are (a) respect for dignity, (b) non-discrimination, (c) accessibility, (d) liberty and security, (e) participation in political and public life, (f) the right to education, work, and health, and (g) protection from exploitation, abuse, and violence. As with the rights of the child, this Convention also has a Committee on the Rights of Persons with Disabilities, which monitors the implementation of measures to eliminate discrimination and promote the rights of persons with disabilities by States Parties.

Since the nature of the project is inextricably linked to security, it should also be highlighted as a fundamental human right. Article 3 of the Universal Declaration of Human Rights (UN, 1948) states that "everyone has the right to security of his or her person", a right that appears to be related to the right to life and liberty. The right to security also can be found in the International Covenant on Civil and Political Rights (UN, 1966a, art. 9.1), in the European Convention on Human Rights (Council of Europe, 1950, art. 5) and in the Charter of Fundamental Rights of the European Union (European Union, 2016a, art. 6). In all three cases, security is linked to the right to life and liberty, underlining its centrality and importance in the human rights regime.

Privacy is a right safeguarded by international agreements, as stated in Article 12 of the Universal Declaration of Human Rights (UN, 1948), Article 17 of the International Covenant on Civil and Political Rights (UN, 1966a) and Article 8 of the European Convention on Human Rights (Council of Europe, 1950). Regulations implemented by the EU have notably emphasised this sphere, establishing a model for worldwide norms. The European Union's Charter of Fundamental Rights (EU, 2016a) recognises the significance of private and family life in Article 7, stating that "Everyone is entitled to the respect of their private and family life, home, and communications". Similarly, Article 8 on data protection emphasises that individuals have the right for their personal data to be handled confidentially. Information must be processed ethically for specific purposes based on the data subject's consent or other legal grounds. Moreover, individuals have the right to access and rectify their accumulated data. The General Data Protection Regulation (GDPR) (European Parliament & Council of the European Union, 2016) holds the most considerable influence in this domain. The GDPR establishes a framework to protect personal data, enhancing individuals' control over their information, while placing significant responsibilities on organisations that manage such data. The main aim of the GDPR is to improve transparency, security, and accountability in the handling of personal information.

While it is true that the European Convention on Human Rights (Council of Europe, 1950) mentions that "Everyone has the right to respect for his private and family life, his home and his correspondence", this interference is permissible when it is “authorized by law and constitutes a necessary measure for national security, public safety, the country's economic well-being, the prevention of disorder or crime, the protection of health or morals, or the protection of the rights and freedoms of others." With respect to the correlation between rights and exceptional situations, Article 15 of the European Convention on Human Rights states that “in times of war or other emergencies that threaten the nation's safety, a High Contracting Party may take measures derogating from its obligations under this Convention, as long as such measures are strictly necessary and not in violation of international law.” Therefore, one must strike a balance between human rights and the unique characteristics of exceptional situations, such as natural or man-made disasters. This does not excuse states from their obligations to uphold human rights, but it does provide a degree of flexibility that can be applied - always in a proportionate manner - to situations where the security of the nation is threatened (for further details see section 2.4).

Freedom of thought is also a strongly protected right in the human rights regime. It is enshrined in Article 18 of the Universal Declaration of Human Rights (UN, 1948): "Everyone has the right to freedom of thought, conscience and religion; this right includes freedom to change his religion or belief, and freedom, either alone or in community with others and in public or private, to manifest his religion or belief in teaching, practice, worship and observance". The same right is also set up in Art. 18 of the International Covenant on Civil and Political Rights (UN, 1966a), Art. 10 of the Charter of Fundamental Rights of the European Union (EU, 2016a) and Art. 9 of the European Convention on Human Rights (Council of Europe, 1950). In line with the above-mentioned right, we can also refer to freedom of opinion and expression, a right that includes the freedom to hold opinions without interference and to seek, receive and impart ideas through any media and regardless of frontiers. This right is enshrined in Art. 19 of the Universal Declaration of Human Rights (UN, 1948), Art. 19 of the International Covenant on Civil and Political Rights (UN, 1966a), Art. 11 of the Charter of Fundamental Rights of the European Union (EU, 2016a) and Art. 10 of the European Convention on Human Rights (Council of Europe, 1950). Finally, Art. 20 of the Universal Declaration of Human Rights (UN, 1948) states that "Everyone has the right to freedom of peaceful assembly and of association". The right to freedom of assembly and association is also recognised in Articles 21 and 22 of the International Covenant on Civil and Political Rights (UN, 1966a), in Art. 12 of the Charter of Fundamental Rights of the European Union (EU, 2016a) and Art. 11 of the European Convention on Human Rights (Council of Europe, 1950).

In accordance with the aforementioned rights, it is also crucial to emphasise the significance of safeguarding consumer interests and welfare. The PANTHEON project's goal is to design a disaster management assistance tool that can be used for various phases of the disaster management cycle. Therefore, it is essential to ensure basic product quality standards to meet consumer rights. This is outlined in Article 12 of the Treaty on the Functioning of the European Union (EU, 2016c), which declares that "Consumer protection requirements shall be taken into account in defining and implementing other Union policies and activities". Consumer rights are also stated in the Charter of Fundamental Rights of the European Union (EU, 2016a), and furthermore, legislation has been established. One example is the Consumer Rights Directive (2011/83/EU), which aims to strengthen and coordinate consumer rights across the Union, promote clarity, protect against unfair commercial practices, and ensure a period of reflection for consumers in off-premises and distance contracts. The Product Safety Regulation (EU Regulation 765/2008) is a pivotal part of this regime with the objective of safeguarding the safety and quality of products placed on the European market. This regulation imposes stringent accreditation mandates on conformity assessment entities, enhances market supervision efficiency, and ascertains Member States collaborate in implementing these rules. Finally, it is noteworthy that EU Regulation 2017/2394, the Regulation on Consumer Protection Cooperation, aims to enhance collaboration between national authorities accountable for enforcing consumer protection laws in the European Union.

Finally, there is the issue of climate change. The environment has become increasingly important worldwide in recent decades, particularly in relation to natural disasters. While it is not always possible to establish climate change as an independent variable in all such disasters, it is widely agreed to be an intervening variable. Similarly, while the effects cannot be attributed solely to human activities (Stern & Kaufmann, 2014), it is undeniable that these activities have had an impact. This is outlined in the most recent report from the Intergovernmental Panel on Climate Change (IPCC, 2023), which asserts that "Human activities, principally through emissions of greenhouse gases, have unequivocally caused global warming, with global surface temperature reaching 1.1°C above 1850–1900 in 2011–2020. Global greenhouse gas emissions have continued to increase, with unequal historical and ongoing contributions arising from unsustainable energy use, land use and land-use change, lifestyles and patterns of consumption and production across regions, between and within countries, and among individuals."

In 1997, the Kyoto Protocol (UN, 1997) was established as the first binding agreement to reduce greenhouse gas emissions. It has since been superseded by the Paris Agreement (UN, 2015), which is widely regarded as the most significant international treaty related to climate change. Adopted during the United Nations Framework Convention on Climate Change, this agreement provides a framework for limiting global warming to below 2 degrees Celsius. At the European level, Article 11 of the Treaty on the Functioning of the European Union (EU, 2016b) stipulates that environmental protection requirements should be integrated into the meaning and execution of the Union's policies and activities, especially for the advancement of sustainable development. Additionally, Article 37 of the Charter of Fundamental Rights of the European Union (EU, 2016a) refers to environmental protection. Similarly, the European Union nations have endorsed the Renewable Energy Directive (EU 2018/2001) to boost its utilization, the Energy Efficiency Directive (EU 2023/955) to enhance energy efficiency, and the EU Emissions Trading System (EU ETS) that aims to foster the reduction of greenhouse gas emissions.

This is not an exhaustive examination of the human rights regime. We aim to emphasise the human rights pertinent to the development of the PANTHEON system/platform and briefly mention the relevant legal texts. The project hinges on equality, non-discrimination, security, privacy, freedom of thought, freedom of opinion and expression, freedom of assembly and association, consumer protection and environmental respect, which are essential components for the product to be developed. The relationship between the main human rights and the texts recognising them is outlined in Table 1.

*Table 1: Main Human Rights instruments*

| Specific right | Universal Declaration of Human Rights | International Covenant on Civil and Political Rights | Convention for the Protection of Human Rights and Fundamental Freedoms | Charter of Fundamental Rights of the European Union |
|---|---|---|---|---|
| **Equality** | Art. 1 | Art. 26 | Protocol nº 12 | Art. 20 |
| **Non-discrimination** | Art. 7 | Art. 26 | Art. 14 | Art. 21 |
| **Security** | Art. 3 | Art. 9 | Art. 5 | Art. 6 |
| **Privacy** | Art. 12 | Art. 17 | Art. 8 | Art. 7 |
| **Freedom of though** | Art. 18 | Art. 18 | Art. 9 | Art. 10 |
| **Freedom of opinion and expression** | Art. 19 | Art. 19 | Art. 10 | Art. 11 |
| **Freedom of assembly and association** | Art. 20 | Art. 20, Art. 21 | Art. 11 | Art. 12 |

## 2.2. HUMAN RIGHTS IN DIGITAL AGE

Information and Communication Technologies (ICT) have radically changed the way we live, work, and interact. While these innovations offer countless benefits, they also pose significant challenges to the human rights regime. This was acknowledged as early as the 1748th plenary session of the United Nations General Assembly on 19 December 1968, which recognised "that recent scientific discoveries and technological advances, while opening up vast prospects for economic, social and cultural progress, may nevertheless endanger the rights and freedoms of individuals and peoples and therefore require constant vigilance" (UN, 1968). However, the application of human rights to technology still has a long way to go. As Perry & Roda (2017, p. 7) write, technological advances do not require new human rights, but rather the extension of existing human rights to new technologies. In the same vein, Coccoli (2017, p. 226) mentions the need to reinterpret human rights considering new technologies. This is precisely why David Wright (2013, cited in Perry & Roda, 2017, p. 80) recommends conducting an impact assessment that integrates human rights into the design of digital technologies.

Technologies such as Artificial Intelligence (AI) and Machine Learning (ML) present a dichotomy in terms of equality, because while they have the potential to reduce inequalities by facilitating access to information and opportunities, they also risk perpetuating and widening existing gaps due to inherent biases in algorithms and lack of equitable access to these innovations. On the one hand, there are differences between people with effective access to technology and those with limited or no access. There are two main components of digital inequality: a geographical one, which refers to the urban-rural and/or developed-underdeveloped cleavage; and a social one, which includes digital literacy, access by vulnerable groups and language barriers, among others (Coccoli, 2017, p. 242). In this regard, a report by the United Nations Secretary-General stated that everyone should have the opportunity to improve their well-being through technological innovation, not only in terms of physical accessibility and digital literacy development, but also by ensuring that the design of technology respects the needs of all people, including vulnerable groups, and addresses intersectionality, social norms and language barriers, among others (UN, 2020, pp. 10-11). On the other hand, when technologies such as AI recommend certain decisions, these may be unfair to certain individuals or groups due to biases. Such biased decisions can have serious consequences in terms of freedom and access to resources, as they may deny certain people a job, a bank loan or even unjustly imprison someone (Coeckelbergh, 2020, pp. 125-127). It is therefore essential to promote transparency in algorithms and diversity in technological development to avoid reproducing stereotypes and discrimination (for more on AI, see section 2.3, section 3.1.4 and D10.3).

In the field of security, the use of advanced technologies can enhance the ability of states to prevent and respond to threats, but their inappropriate use also poses significant human rights risks. AI, big data analytics and cybersecurity have certainly changed the way we approach security. On the one hand, the same technologies that enable greater freedoms can also be used to restrict them by filtering and/or blocking certain content and even denying access to certain technologies (Coccoli, 2017, p. 228). The quintessential example is the Chinese government's control of access to content that crosses its digital borders through the so-called 'Great Firewall'. The purpose of this type of censorship is ostensibly to sever social ties when there is a risk or evidence of collective movement, thereby reducing the likelihood of collective action (Perry & Roda, 2017, p. 103). Chinese law even contemplates shutting down the internet in certain regions under the pretext of protecting national security and public order (ibid, p. 107). Although China is an extreme case, Western societies - especially since the terrorist attacks of 11 September 2001 - have not been unaffected by the impact of new technologies on surveillance practices and have also experienced their use in devices such as CCTV, video cameras, satellite surveillance, fingerprinting and other forms of biometric monitoring

(Hiranandani, 2011, pp. 1091-1093). For example, after the first attacks on the satirical magazine Charlie Hebdo, the French parliament pushed through a surveillance law that allowed the interception of all online data and communications generated by individuals with alleged suspicious behaviour (Perry & Roda, 2017, p. 109). While privacy is the main victim of such practices, they also affect other fundamental rights such as freedom of opinion, expression, assembly, and association (UN, 2014, p. 5). In particular, public spaces provide the basis for informal socialisation and civic life, and the threat of surveillance can sometimes force citizens to protect their privacy, thereby reducing social interaction and the value of public spaces (Hiranandani, 2011, p. 1099). In addition, surveillance is often targeted at specific individuals or groups, thereby fostering discrimination. Indeed, in the aftermath of the 9/11 attacks, immigrants, people of Arab descent, people from the Middle East and South Asia, and international students were the most targeted groups for surveillance (ibid, p. 1096). The digital age thus has the capacity to affect the democratic culture of society. However, new technological structures also increase the opportunities for individuals and collectives to participate in their respective cultures, thus fostering both individual and collective identities.

Apart from using security as a pretext to violate human rights such as privacy, freedom of opinion, expression and assembly, the digital age has also brought new security challenges. For example, while the emergence of smart cities offers a great opportunity to improve the well-being of citizens, including their security, it also poses significant security challenges (for more on Smart Cities see section 3.1.1). In this regard, there is considerable consensus that Smart Cities are being designed and implemented with few security guarantees (Asif et al., 2022, p. 3; Al-Turjman et al., 2022, p. 8; Cerrudo, 2015 cited in Kitchin & Dodge, 2019, p. 50; Cui et al., 2018, p. 46137) and are also highly vulnerable to cyberattacks (Alamer & Almaiah, 2021, p. 719). The emergence of new technologies and their consolidation in society has created new risks and vulnerabilities, which are exploited by cybercriminals. In this context, security necessarily becomes a fundamental challenge to be addressed, and designers of technological products must take it into account and provide consumers with robust and sustainable solutions (Perry & Roda, 2017, p. 148).

Regarding privacy, some authors refer to it as the digital right par excellence (Coccoli, 2017, p. 238). In context of the rapidly progressing digitalisation of society citizens are not fully aware of the scale of data collection, which is almost ubiquitous (Perry & Roda, 2017, p.71). Most people expect that their individual behaviour will not be observed, monitored, or recorded without their consent (Hiranandani, 2011, p. 1092), so privacy violations can be considered as a form of aggression against personal identity. A person can be seen as his or her information and, consequently, whatever is done to his or her information is done to that person and not to his or her property (ibid, p. 1098). Furthermore, privacy also affects other rights, such as freedom of association and expression, and can facilitate different types of discrimination (Perry & Roda, 2017, p. 138).

Finally, there is the issue of climate change and sustainability. Many of today's technological systems require the extraction of rare minerals. Apart from the fact that the extraction of these minerals often leads to human rights violations - especially in developing countries and vulnerable groups - the scarcity of these resources and the difficulties in supplying an exponentially growing technology market must be taken into account (ibid, p. 13). This, of course, has a direct impact on the sustainability of such products. On the other hand, some technological systems require high energy consumption to function (ibid, p. 135) and designers and developers should take this into account.

In the previous section, we made a compilation of the most relevant human rights, typified at both international and European level. Following on from this, the aim of this section is to show how the digitalisation of society affects these rights. Bias and discriminatory algorithms certainly pose a risk to equality, as do unequal access to technology and different skills in its use. Technology has also been an element that has enabled greater surveillance of the population, affecting rights such as privacy, but also freedom of expression and assembly. Moreover, many of these systems have been developed without sufficient consideration of their security, creating risks and vulnerabilities that, if exploited, could lead to serious human rights violations. Finally, the scarcity of raw materials needed to manufacture many devices and the high demand from the technology sector pose sustainability challenges that need to be seriously addressed. In addition, although many energy-efficient devices are already in use, energy efficiency is also an element to be considered in the design and development of sustainable and environmentally friendly technological systems.

Given the impact that new technologies can have on the human rights regime, the European Union has equipped itself with specific legislation to address emerging challenges. One of these tools is the EU Network and Information Security Directive (EU NIS Directive). It was adopted in 2016 with the aim of improving cybersecurity in the EU and increasing cooperation and protection of critical infrastructure such as energy, transport, financial services, and digital service providers. In January 2023, Directive 2022/2555 (NIS2) replaced the previous directive, rendering it legally ineffective. The new directive requires member states to adopt national cybersecurity strategies and establish cybersecurity crisis management authorities. It also sets out cybersecurity risk management measures and incident reporting obligations. The text of the Directive is binding on Member States, with monitoring mechanisms to ensure that they implement the provisions in their national legislation.

Similarly, in December 2022, the President of the Council of the European Union, the European Parliament and the Commission signed the Declaration on Digital Rights and Principles for the Digital Decade, which aims to help the EU achieve its digital transformation goals while respecting the values on which the Union is founded. Chapter I underlines the importance of putting people at the centre of the digital transformation to ensure that the process benefits everyone and that individual rights are respected online. Chapter II mentions guarantees for all EU citizens to have access to high quality internet and adequate training to develop digital skills, thus underlining the will to promote equality.

Chapter III refers to the right to non-discrimination by mentioning that all algorithmic systems used in the EU must be trained with accurate and rigorous data to avoid bias and discrimination. The same chapter also mentions the protection of users' and consumers' rights and calls for increasing their trust in new technologies and enabling them to make informed choices in this regard. Chapter IV reiterates that all EU citizens have the right to secure, diverse and reliable access to the digital environment. Diversity of content must be preserved to promote the democratic principles on which the EU is founded, but at the same time citizens must be protected from phenomena such as disinformation, manipulated information and other content that could cause them harm. Reference is also made to freedom of opinion, expression, assembly, and association, which must be safeguarded in the digital context.

Chapter V deals with safety and privacy. On the one hand, the EU must ensure that only safe and compliant products enter its market. On the other hand, it is also responsible for ensuring control over its citizens' personal and non-personal data and the privacy of their communications. EU citizens must be protected from cyber risks and threats such as mass surveillance and illegal interception of communications. Finally, Chapter VI is dedicated to sustainability and sets out the EU's willingness to develop and deploy sustainable technologies with minimal social and environmental impact.

Therefore, in view of the strong impact of technology on human rights and the increasing efforts to regulate it, its inclusion and integration into the PANTHEON system will not only be a necessity, but also an added value both for its consumers and for society in general.

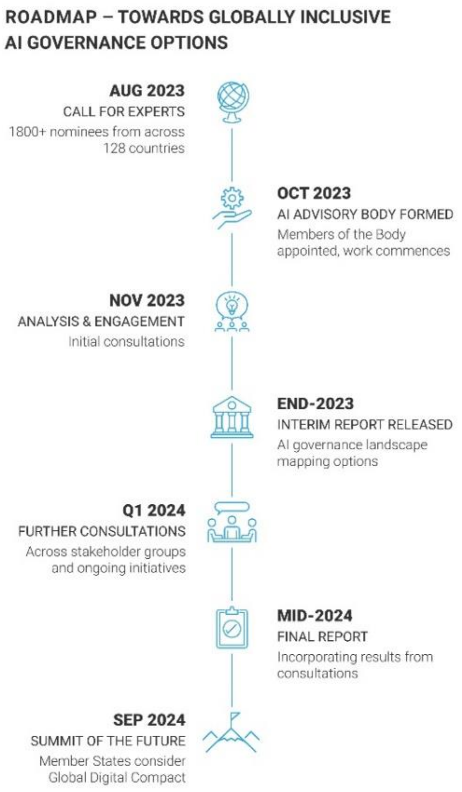
## 2.3. ARTIFICIAL INTELLIGENCE

In the previous section, we highlighted and demonstrated the impact that new technologies can have on human rights. Artificial Intelligence (AI) is arguably the technology that has most revolutionised our society in recent years, and its future development is still uncertain. Indeed, the speed and scope of AI is unprecedented and the gap between this technology and its governance is wide and increasing (UN, 2023a). Broadly speaking, we could define AI systems as those that have the ability to automatically process data and information in a human-like manner, including aspects such as reasoning, learning, perception, prediction or planning (UN, 2022). In particular, one of the aspects that most concerns the PANTHEON project is the ability to make predictions and recommendations, thus influencing real and/or virtual environments with clearly defined goals by humans (OECD, 2019). On the other hand, it is also worth clarifying the concept of the "AI lifecycle", which consists of the process of (1) design, (2) verification and validation, (3) deployment, and (4) operation and monitoring (ibid.).

AI has the potential to improve human well-being, contribute to sustainable global economic activity, increase innovation and productivity, and respond to key global challenges (OECD, 2019). It seems clear that AI systems are capable of advancing societal progress and, in particular, improving crisis forecasting and management (UN, 2023b). However, despite the opportunities for it to be a positive force, it also brings with it a number of challenges and risks arising from its potential malicious and malevolent use, as well as the inequalities it could create or exacerbate (UN, 2022). Technology has also been assessed in the context of democracy and human rights, concluding that while it can have a positive impact on these, there is a risk of undermining democratic principles and violating fundamental rights (UN, 2023b; OECD, 2019). In particular, reference was made to misinformation and manipulation of public opinion by generative AI, but also to discrimination caused by algorithms with biases and prejudices, security and cybersecurity risks, and threats to privacy (OECD, 2023, pp. 13-14). In this regard, the UN Secretary-General has identified three areas where smart technologies could pose a serious risk to the human rights regime. On the one hand, in terms of threats that are already materialising today, there is the marketing of AI models with little regard for security, which exponentially increases the risk of their exploitation by criminals and even terrorists. On the other hand, in terms of long-term consequences, he speaks of disruptions in labour markets and the economy, the loss of cultural diversity as a result of discriminatory algorithms, disinformation, manipulation and mass surveillance. Finally, it points to the exacerbation of existing inequalities, as these types of technologies are limited to a few countries and companies that have the technical and economic capacity to benefit from them (UN, 2023a).

Considering some of the aspects that could have a negative impact on people and their rights, the need to establish governance and regulation for this type of technology is clear. While it is true that there are difficulties and challenges in legislating for AI (OECD, 2023, pp. 22-23), the latter does not require new principles and should be based on the United Nations Charter and the Universal Declaration of Human Rights (United Nations, 2023a). The Organisation for Economic Co-operation and Development has also highlighted the existence of legal frameworks within which it would be possible to frame AI, such as the human rights regime, consumer protection, personal data protection, intellectual property rights, etc., while recognising the need to create new frameworks to address the challenges associated with AI (OECD, 2019).

At the international level, the UN Secretary-General has already established the High-Level Advisory Body on Artificial Intelligence in 2023 to prepare an analysis and make advanced recommendations for the international governance of AI. The panel is composed of government, private sector and civil society actors who are experts in the relevant disciplines. The aim is to offer diverse perspectives and options on how to govern AI for the public good, while aligning this technology with human rights and the Sustainable Development Goals. The group is expected to publish an interim report by the end of 2023 and a final version by mid-2024.


ROADMAP – TOWARDS GLOBALLY INCLUSIVE AI GOVERNANCE OPTIONS
AUG 2023
CALL FOR EXPERTS
1800+ nominees from across 128 countries
OCT 2023
AI ADVISORY BODY FORMED
Members of the Body appointed, work commences
NOV 2023
ANALYSIS & ENGAGEMENT
Initial consultations
END-2023
INTERIM REPORT RELEASED
AI governance landscape mapping options
Q1 2024
FURTHER CONSULTATIONS
Across stakeholder groups and ongoing initiatives
MID-2024
FINAL REPORT
Incorporating results from consultations
SEP 2024
SUMMIT OF THE FUTURE
Member States consider Global Digital Compact

*Source: https://www.un.org/techenvoy/ai-advisory-body*

Figure 1 – Chronology of the UN's actions on AI

Within the United Nations, the United Nations Educational, Scientific and Cultural Organisation (UNESCO, 2022, pp. 20-23) has developed a set of principles and recommendations for ethical AI. The principles set out are as follows:

- *Do no harm*: Throughout its life cycle, AI should avoid any individual and/or collective harm, including social, cultural, economic, and political harm. The design, development and deployment of an AI system must respect and protect human rights and fundamental freedoms.
- *Purpose, necessity, and proportionality*: the use of AI must be justified, appropriate and not go beyond what is strictly and fundamentally necessary.
- *Safety and security*: any security-related risks, including cybersecurity, must be identified, addressed, and mitigated to prevent and limit actual or potential harm.
- *Fairness and non-discrimination*: The benefits of AI must be distributed in a fair and equitable manner. In addition, people designing, developing, and implementing such technologies should prevent bias, discrimination, and stigmatisation by algorithms.
- *Sustainability*: Human, social, cultural, political, economic, and environmental impacts must be assessed on an ongoing basis.
- *Right to privacy and data protection*: Privacy must be protected throughout the lifecycle of AI, using appropriate personal data protection measures.

- *Autonomy and human control*: AI design should be human-centered and human control over the technology should be ensured.
- *Transparency and accountability*: There must be maximum transparency in the decisions made by AI systems.
- *Responsibility and accountability*: There must be appropriate human control of AI, analysis and assessment of its impact, and due diligence mechanisms.
- *Awareness and literacy*: All people need to be aware of the characteristics and capabilities of AI and have the skills to use it.
- *Inclusion and participation*: the design, development and implementation of AI systems should be interdisciplinary and participatory, involving as many stakeholders as possible.

On the other hand, the Organisation for Economic Co-operation and Development (OECD, 2019) developed the first intergovernmental standard for AI-related policies. The principles set out are:

- *Inclusive growth, sustainable development, and well-being:* AI should pursue goals and outcomes that benefit people and the planet. They should enhance - not replace - human capabilities and inclusiveness, reduce social, economic and gender inequalities, and protect the natural environment.
- *Human-centered values and justice*: AI systems must respect the rule of law, human rights, and democratic values throughout their life cycle.
- *Transparency and accountability*: general understanding of AI should be promoted, stakeholders should be made aware that they are working with AI, the results of AI should be understandable to those affected, and complaints, claims and grievance processes should be facilitated for those adversely affected.
- *Robustness and security*: AI systems must be robust and secure throughout their lifecycle. There must also be traceability of the data, processes and decisions made throughout their lifecycle. Risk management must be systematically and continuously applied at every stage of the system to address system risks, including privacy and digital security.
- *Accountability*: AI actors must be held accountable for the operation of AI systems, in particular regarding the above principles.

At the regional level, one of the most relevant legislative proposals currently underway is the European Union (EU) AI Act. Already in April 2021, the Commission published a proposal to regulate artificial intelligence in the EU. In December 2022, the EU Council adopted its common position on the legislative proposal and in June 2023, the European Parliament approved the text with 499 votes in favour, 28 against and 93 abstentions. The proposal will thus become law if the EU Council and the European Parliament agree on a common version of the text. The regulation is part of the EU's digital strategy and aims to create better conditions for the development and use of this innovative technology. In particular, it aims to ensure that AI systems entering the EU market are safe and respect existing laws in the region. It also intends to provide legal certainty to facilitate investment in such systems and to ensure the enforcement of European laws on fundamental rights and safety requirements applicable to AI (Art. 1 AI Act). The rules set out in the proposed legislation will impose obligations on providers and users depending on the level of risk posed by each AI system. This legislation has been discussed in more detail in D10.3 (Shtefchyk & Vallès, 2023).

In terms of national legislation, several countries have already taken the initiative to regulate these technologies. The Canadian Parliament proposed the Artificial Intelligence and Data Act (AIDA) in June 2022, which is also risk-based and aims to promote the responsible design, development, and use of AI in the Canadian private sector. France has already developed a national AI strategy, as have Germany, Italy, and

Japan. The UK has adopted the AI Regulation White Paper, and in the case of the US, the situation has been addressed through voluntary commitments by leading AI companies to promote safe and transparent AI development (OECD, 2023, pp. 32-36). However, as our pilot areas are Vienna and Athens, it is important to note the initiatives taken by Austria and Greece in this regard.

On the Greek side, the most relevant legislation in this regard is Law 4961/2022, which regulates the use and exploitation of advanced technologies that have a significant social and economic impact. The law establishes vertical and horizontal obligations for public institutions, private entities and individuals that produce, distribute and use any of the regulated technologies (Mitsios, 2022). Its objectives include the implementation of a national AI strategy, the completion of the legal framework on cybersecurity and the promotion of new measures for the transparent and safe operation of IoT devices, unmanned aerial systems (UAS), distributed ledger systems, 3D printing technologies and smart contracts (Argyri et al., 2022). Part A, which runs from Art. 1 to 27, establishes an institutional framework for the operation of AI systems, both in the public and private sectors, under conditions of fairness and security, and strengthens resilience against cyber threats (Mitsios, 2022). Its Art. 11 establishes the Steering Committee on AI, which is responsible for designing and promoting AI policies and actions, formulating national AI priorities, improving the national AI strategy, and taking decisions on its implementation. Art. 13 establishes the executive body of the Steering Committee: The Committee for the Supervision of the National AI Strategy, which is responsible for coordinating the national strategy, ensuring the implementation of the decisions of the Steering Committee and reporting on the progress of implementation. Similarly, Art. 14 establishes an AI Observatory within the Ministry of Digital Governance (Argyri et al., 2022).

While it is true that the implementation of an AI strategy is one of its main tasks, the Act also considers other emerging technologies. Of particular interest is the IoT, as it establishes a series of obligations for manufacturers, importers, distributors, and operators of such technology. In its Art. 32, the law provides for security measures for IoT devices, which must be designed with an adequate level of safety and cybersecurity (ibid).

Austria has actively contributed to the creation of legal and ethical frameworks in the field of AI. At the national level, it established the Austrian Council for Robotics and Artificial Intelligence (ACRAI) in 2017 and developed an Austrian White Paper on AI. At the international level, it has also been very active in the discussions of the European High-Level Expert Group on Artificial Intelligence (HLEG-AI) to develop ethical guidelines for trustworthy AI, as well as in the development of the UNESCO Recommendations on AI mentioned above. Despite their strong involvement in this area, they have not created a specific law at the national level. However, they have initiated a consultation process for an "Artificial Intelligence Mission Austria 2030", which sets out the national strategy for the use of AI and recognises the risks of AI regarding the labour market, democracy, human rights and the restriction of freedoms (Federal Ministry for Climate Protection, Environment, Energy, Mobility, Innovation and Technology, 2021).

The strategy has two salient features. On the one hand, it is a European strategy based on its values, the rule of law and fundamental human rights. On the other hand, it is a broadly participatory strategy, bringing together more than 160 experts from disciplines as diverse as technology, economics, law, natural sciences, social sciences, and education, as well as civil society organisations and Austrian citizens. The objectives of the strategy are:

- Ensure the deployment of AI systems for the public good, in a responsible manner and based on human rights, European values and future European AI regulation.
- Make Austria a leading international player in AI research and investment.

- Securing the competitiveness of Austrian technology.

To achieve this, the strategy identifies fields of action that focus on creating reliable AI and a suitable ecosystem for its design, development, and deployment. It also identifies the most relevant application areas, including climate change, energy systems, sustainable mobility, agriculture, and smart cities. Finally, it should be noted that this is not a rigid or static strategy, but a living document. Of course, the strategy is open to adjustments and clarifications and aims to involve as many stakeholders as possible in its development.

## 2.4. LEGAL FRAMEWORK IN EMERGENCY SITUATIONS

So far, we have seen the importance of certain human rights, the main legal texts in which they are enshrined, and the impact of the emergence of new technologies, in particular AI systems, on them. However, as mentioned in section 2.1, rights are not absolute and there are situations in which they can be limited by public authorities, as long as a greater public good is at stake and this is done responsibly and proportionately. Such measures are deeply rooted in democratic regimes and are necessary to preserve their fundamental values, which is why they are regulated in the above-mentioned texts.

For example, the International Covenant on Civil and Political Rights, which typifies the right to freedom of thought, expression, assembly, and association, considers the limitation of each right in exceptional situations. Art. 18.3 states that "freedom to manifest one's religion or beliefs may be subject only to such limitations as are prescribed by law and are necessary to protect public safety, order, health or morals or the fundamental rights and freedoms of others". Similarly, Article 19(3) on freedom of expression states that "the exercise of the rights set forth in paragraph 2 of this article entails special duties and responsibilities. It may therefore be subject to certain restrictions, but only such as are provided by law and are necessary: (a) for the respect of the rights or reputations of others; (b) for the protection of national security or public order, or of public health or morals". The same is true of Articles 21 and 22, which regulate the rights of assembly and association and provide for their restriction only in situations established by law and which seriously threaten the interests of national security or public safety, public order, public health and the protection of the rights and freedoms of others. Likewise, the European Convention on Human Rights, in its Articles 9, 10 and 11 on freedom of thought, expression, assembly and association, contemplates the restriction of these rights in the same cases as above.

As regards the regional level, Article 52 of the Charter of Fundamental Rights of the European Union states that "any limitation on the exercise of the rights and freedoms recognised by this Charter must be provided for by law and respect the essence of those rights and freedoms. Subject to the principle of proportionality, such limitations may be imposed only if they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others". Furthermore, if we focus on specific legislation, the EU General Data Protection Regulation also provides for restrictions in its Art. 23. In particular, it mentions that the Union and the Member States to which the controller or processor belongs may restrict the rights set out between Art. 12 and 22 (information and communication rights, right of access to personal data, right of rectification and erasure, right to restrict processing, right of data portability and right to object, including in the case of automated processing), Art. 34 (notification of a personal data breach to the data subject) and Art. 5, as long as the restrictions respect the essence of fundamental rights and freedoms and are necessary and proportionate in democratic societies in order to protect:

a) National security
b) Defence
c) Public security

d) The prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including the protection against and prevention of threats to public security
e) Other important objectives of general public interest of the Union or of the Member States, in particular an important economic or financial interest, including monetary, budgetary, fiscal, public health and social security matters
f) The protection of the independence of the judiciary and of judicial proceedings
g) The prevention, investigation, detection, and prosecution of breaches of ethics in the regulated professions
h) A supervisory, inspection or regulatory function relating to the exercise of public authority in the cases referred to in points (a) to (e) and (g)
i) The protection of the data subject or of the rights and freedoms of others
j) The enforcement of civil claims

While it is true that the right to privacy and the protection of personal data are subject to the limitations mentioned above, these limitations should include specific provisions on the following issues:

a) The purposes or categories of processing
b) The categories of personal data
c) The scope of the restrictions imposed
d) Safeguards to prevent misuse or unlawful access or transfer
e) The specifications or categories of controllers
f) The storage periods and safeguards applicable having regard to the nature, scope and purposes or categories of processing
g) The risks to the rights and freedoms of the data subject
h) The data subject's right to be informed of the restriction unless this would be detrimental to the purpose of the restriction.

As far as national legislation is concerned, the Greek Constitution (Constitution of Greece, 2001) contains some provisions that limit or restrict certain rights in cases of emergency. One of these provisions is Article 48, which stipulates that "in case of war or mobilization owing to external dangers or an imminent threat against national security, as well as in case of an armed coup aiming to overthrow the democratic regime, the Parliament, issuing a resolution upon a proposal of the Cabinet, puts into effect throughout the State, or in parts thereof the statute on the state of siege, establishes extraordinary courts and suspends the force of the provisions of articles 5 paragraph 4, 6, 8, 9, 11, 12 paragraphs 1 to 4 included, 14, 19, 22 paragraph 3, 23, 96 paragraph 4, and 97, in whole or in part". The rights affected by the state of siege are freedom of movement, the prohibition of arrest without warrant and habeas corpus, the right to privacy, the right to privacy of correspondence and communications, the prohibition of compulsory labour, freedom of association, and the right to freedom of expression and assembly. However, the provision that is most frequently used in emergency situations in Greece is Art. 44, which is independent of Art. 48 and grants legislative powers to the executive in extraordinary circumstances (Kouroutakis, 2019). In particular, the first paragraph of the article states that 'in extraordinary circumstances of urgent and unforeseeable need, the President of the Republic, on the proposal of the Cabinet, may issue acts of legislative content'. This article was the legal basis for dealing with the aftermath of the Thessaloniki earthquake in 1978 and the unprecedented pollution in Athens in 1982. More recently, it has also been used to deal with the economic crisis. It is therefore a tool that can be used in different types of crisis situations and disasters, including natural and man-made ones (Kouroutakis, 2019). Finally, it is also worth mentioning Articles 18 and 22 of the Constitution. Paragraph 3 of Art. 18 states that "the requisition of property for the needs of the armed forces in the event of war or mobilisation, or for the purpose of meeting an immediate social emergency which may endanger public order or health, shall be regulated by special laws". On the other hand, Art. 22(4) mentions that "special laws shall regulate the requisition of personal services in case of war or mobilisation, or to meet the defence needs of the country, or urgent social emergencies resulting from disasters or liable to endanger public health, as well as the contribution of personal work to local government agencies to satisfy local needs".

As far as Austria is concerned, its constitution (Bundes-Verfassungsgesetz, 1920) contains few provisions for dealing with crisis or emergency situations. Moreover, most of them do not refer to situations in which the life of the nation is threatened, but to situations in which the parliament is not available or not functioning, and the main measure will be the transfer of legislative powers. For example, Article 18(3) states: "If the immediate enactment of measures which, according to the Constitution, require a decision  by the National Council, is for the prevention of obvious and irreparable damage to the general public at a time when the National Council is not in session, or if it cannot meet in time, or if it is prevented from meeting by force majeure, the Federal President may, on the proposal of the Federal Government on his and their responsibility, enact these measures by provisional law-amending ordinances.” In this sense, the federal constitutional law allows the Federal President and the state governments to issue provisional ordinances if the federal or state parliaments are incapacitated. These ordinances take the form of administrative laws and can amend or change legislation but are nevertheless provisional measures until the respective parliament can convene (Stöger, 2021).

# 3. PANTHEON'S ETHICAL APPROACH TO TECHNOLOGY

## 3.1. PANTHEON'S CORE TECHNOLOGIES

At the time of writing, the idea and characteristics of PANTHEON are not yet finalised. The objective of WP3 is to analyse the state of the art in terms of technologies that can serve the project, as well as operational procedures for Disaster Risk Management (DRM). It also aims to identify and develop case studies based on the results obtained in WP2. Finally, as far as this report is concerned, WP3 also aims to define some ethical, legal, and social requirements to ensure that the development of the PANTHEON platform is in line with current legal and ethical frameworks. It is important to emphasise the complexity of the latter task, as the possible outcomes of technological innovations are not always foreseeable, and even less so when their functionalities and main characteristics are not yet fully defined. However, this has not prevented a preliminary analysis of the ethical and legal issues that may arise during the design, development, and deployment of the technology, considering the progress made so far.

On the one hand, it is worth highlighting the potential applications identified for PANTHEON, a work carried out in the context of Task 3.2, the result of which can be found in Deliverable 3.2 (Bittner et al., 2023). The following potential applications have therefore been identified (see Table 2):

*Table 2: Potential applications identified for PANTHEON*

| Disaster phase | Application | Description |
|---|---|---|
| **Before a disaster** | Planning and early warning according to simulations | Using models and simulations based on big amounts of valid data to calculate possible evolutions of scenarios. Especially for scenarios for which little experience exists in the region, this could enable the development of emergency plans. Certain structures (e.g., staging areas) could be pre-defined for specific locations. Also, large-scale evacuations that cannot be trained may be planned according to simulations. The system might be used to give a prognosis of best-case and worst-case consequences for different scenarios that can be adapted, or it may serve as an early warning system, constantly monitoring the modelled area. |
| | Training and exercises | Training realistic scenarios is very important for first responders and disaster response, but large-scale exercises are very complex to organize and expensive. Table-top exercises, on the other hand, tend to require high levels of phantasy. A SCDT could fill this gap and even facilitate regular cross-organisational trainings increasing the mutual understanding of other organisation's needs. Models can simulate the evolution of situations and shed light on approaches and priorities that have not been thought of yet. Also, the system may be used to map and develop different scenarios for large-scale exercises. |
| **During a disaster** | Situational picture | During the first phase of an emergency (chaos phase) the system might be helpful to assess the situation on site and help in the early estimation of potential damage. It may also allow to quickly identify surrounding vulnerabilities (critical infrastructure, schools |

| | | |
|---|---|---|
| | | etc.), classify the incident and deliver crucial data for the orientation of emergency services. |
| | Cross-organisational communication | Different organisations have different information available (especially during the early phase of a disaster). If all involved organisations have real-time access to the system, it may serve to share information across organisational boundaries and enhance cross-organisational communication and coordination. |
| **After a disaster** | Documentation and evaluation | Operations may be documented within a SCDT. In other words, a digital twin of operations is designed within the SCDT. This could either happen according to the input during the disaster response and its storage or through the replication afterwards. This use case could facilitate better understanding and transparency of actions taken and support legal security of operations, while providing an ideal foundation for evaluation of operational strategy. |

*Source: Bittner et al., 2023, pp. 69-70.*

On this basis, the PANTHEON consortium agreed during the Task 3.6 workshops that the PANTHEON system/platform will focus on the pre-disaster phase, including both planning and early warning as well as training of relevant disaster actors. This decision was not taken arbitrarily, of course, but taking into account the different perspectives of potential end-users.

On the other hand, potential technologies that could be integrated into the PANTHEON system were also identified. The study of available technologies was carried out as part of Task 3.1 and its results can be found in Deliverable 3.1 (Tsaloukidis et al., 2023).

The main technology to be used by PANTHEON is the Smart City Digital Twin (SCDT) (see Figure 2). To better understand its concept, it is useful to divide it into two parts. On the one hand, although the concept of Smart City is still under debate (Ziosi et al., 2022), we could broadly define it as the use of Information and Communication Technologies (ICT) to improve the services provided in a city and, consequently, the well-being of its citizens (Alamer & Almaiah, 2021, p. 719). To achieve this, smart cities use various devices installed in the city - mostly sensors - to collect data and information related to urban life. The devices are integrated through the Internet of Things and their data and information are analysed by algorithms based on artificial intelligence (Ahmad et al., 2022, p. 2). On the other hand, a digital twin is a virtual representation of an object, system or process created by integrating sensor data, historical data, and simulation models. The SCDT is an advanced application of this technology that allows the modelling and simulation of a complete urban environment of a smart city (Tsaloukidis et al., 2023, p. 17). In terms of disasters, its functionalities allow predicting the dynamics of a city in the event of stressors such as heat waves, hurricanes, or epidemics. SCDTs have the ability to visualise a situation that is very close to reality, thus enabling informed decision making. It is therefore a very useful tool for the different applications mentioned above, in particular for those chosen for the project: disaster preparedness and planning.

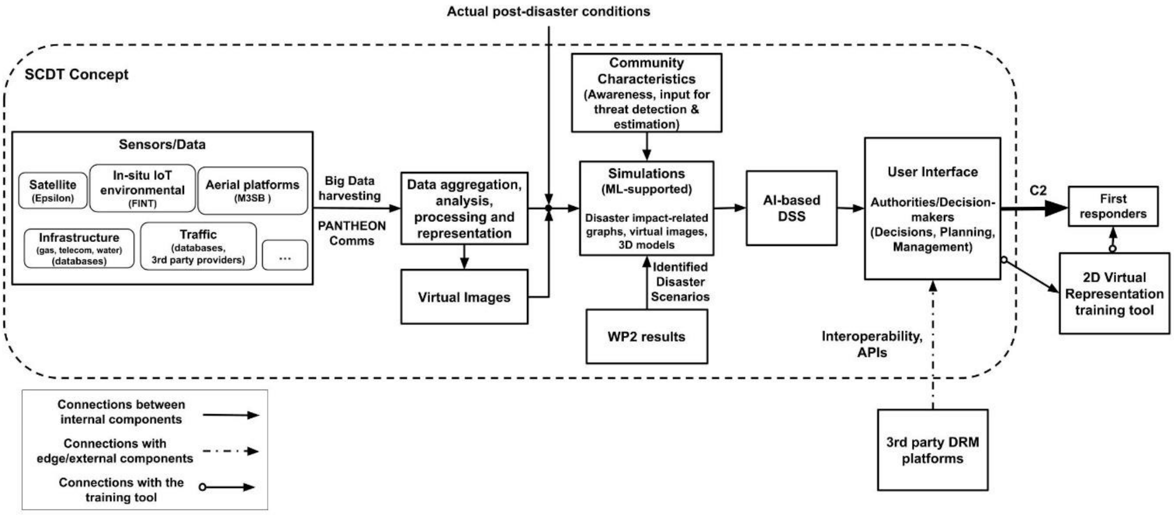
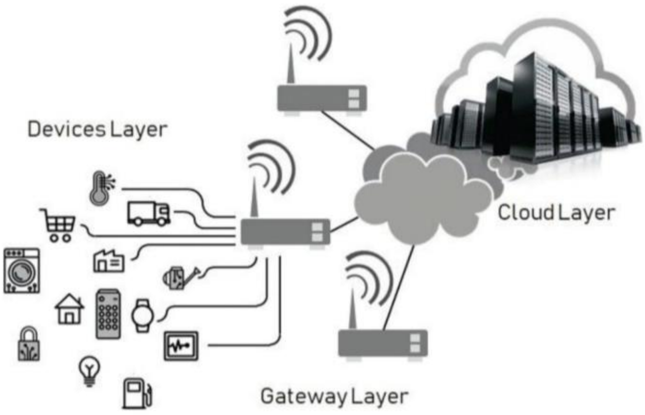

Actual post-disaster conditions
SCDT Concept
Sensors/Data
Satellite
(Epsilon)
In-situ IoT environmental
(FINT)
Aerial platforms
(M3SB )
Infrastructure
(gas, telecom, water)
(databases)
Traffic
(databases, 3rd party providers)
...
Big Data harvesting
PANTHEON Comms
Data aggregation, analysis, processing and representation
Virtual Images
Community Characteristics
(Awareness, input for threat detection & estimation)
Simulations
(ML-supported)
Disaster impact-related graphs, virtual images, 3D models
Identified Disaster Scenarios
WP2 results
AI-based DSS
User Interface
Authorities/Decision-makers
(Decisions, Planning, Management)
Interoperability, APIs
3rd party DRM platforms
C2
First responders
2D Virtual Representation training tool
Connections between internal components
Connections with edge/external components
Connections with the training tool

Figure 2 – PANTHEON approach to SCDT architecture

The digital twin concept is closely related to the development of Internet of Things (IoT) applications. The idea of IoT is fundamentally based on providing connectivity to physical devices via the Internet, mainly as a method of accessing and transferring data, but also for controlling and signalling devices (Bonney et al., 2022, p. 2). An IoT system is built through a layered architecture, which is the organisation of the different elements of the system to meet the needs of the organisation, the environment and end users, as well as operational constraints. While it is true that there is no universal IoT architecture, there is broad consensus on the model of three layers: cloud, gateway, and devices (Wheeler et al., 2020, pp. 47-52) (see Figure 3).



*Source: Wheeler et al., 2020, p.8*

Figure 3 - Three-layered model of IoT

The **cloud layer** typically contains the analytics engines, AI and web applications that are used to process the data collected by an IoT system and subsequently present it in the form of system models for its operators or end users. The most notable feature of this layer is its ability to reorganise massive computing sources to provide data services (storage, computation, or communication) to multiple parties and to share these resources on demand.

The **gateway layer** is the middle layer of the architecture and is responsible for data management, communication, and device control. In fact, gateways can be thought of as processing and communication hubs that provide services directly to the device layer.

Finally, the **device layer** consists of the edge nodes that interact with the physical world. In general, devices can be divided into three types: tags, sensors, and actuators. Tags contain the device identifier and transmit it when required. Sensors translate the physical properties they sense from an environment into digital representations. The most common sensors are temperature, humidity, barometric pressure, and light sensors. Actuators are responsible for interacting with the physical system.

However, as mentioned above, there is no universal approach to systems, so each project has to design its own layered model depending on its needs and the functionality required for its development. Although the final architecture of the system has not yet been determined, preliminary models have been designed (see Figure 4). It should be stressed that this is not the final model, so there is a possibility that some changes will be made to the design soon to better adapt it to the needs of the end users.

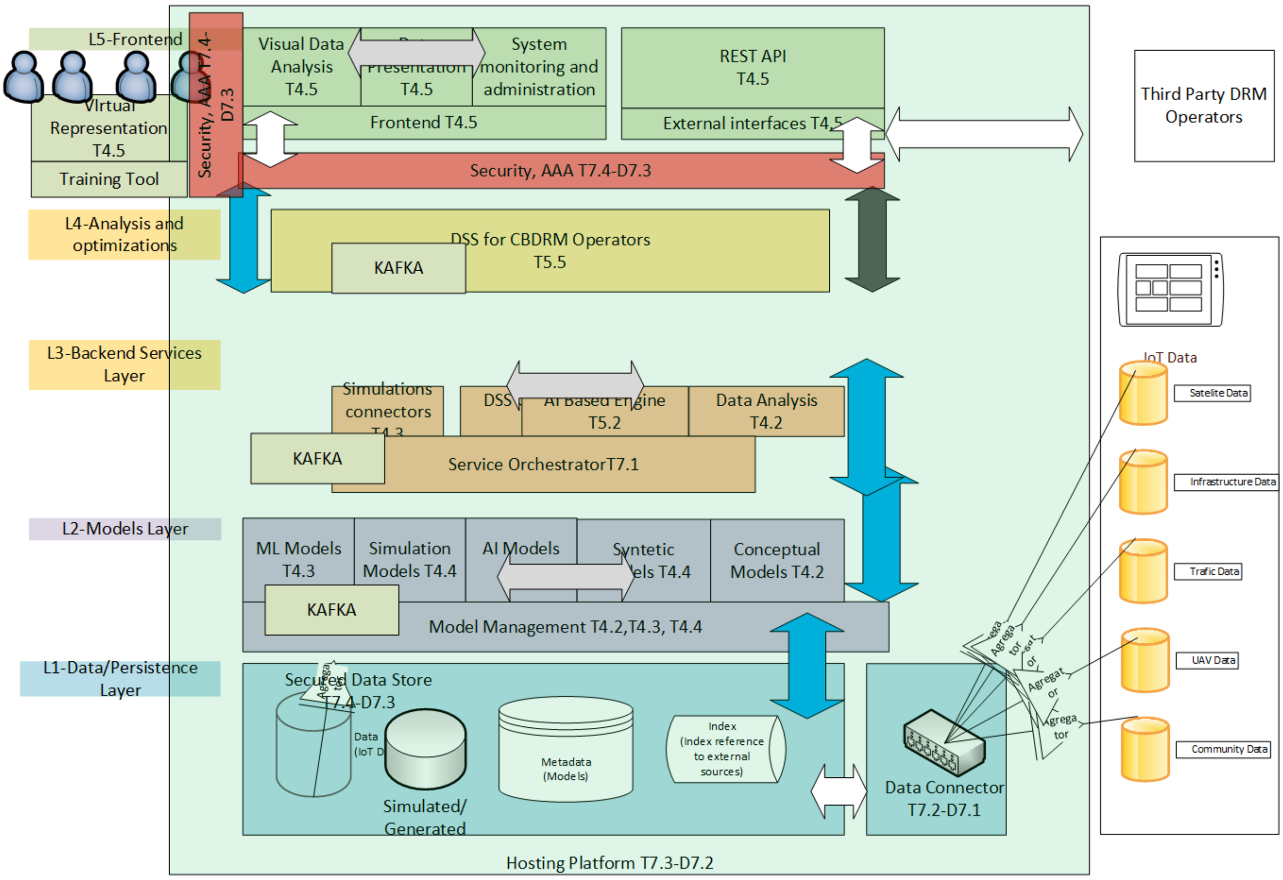


Figure 4 – PANTHEON logical architecture

As can be seen in Figure 4, there are several layers where AI is involved. AI-based modelling refers to the use of AI techniques and algorithms to build and develop models that can make predictions, generate outcomes or perform specific tasks. These models can be built using either machine learning (ML) or deep learning (DL). ML models are trained on historical data to identify patterns and make predictions. Typically, ML-trained systems have the ability to learn and improve through previous experience, and their success depends heavily on the availability of large amounts of data of significant quality, computational power, and the correct design, training and refinement of the model (Tsaloukidis et al., 2023, pp. 23-27).

Another technology under consideration is Decision Support Systems (DSS), which have an interactive nature and are designed to assist those responsible for making a relevant decision. Therefore, their main objective is to improve the effectiveness of the decision-making process through different types of data and the relationships between them. The technology includes a database and a management system to store and make available the information needed to decide. It also consists of a knowledge management system that shows the existing relationships between very complex data and alternative solutions based on problem solving. In parallel, there is a model management system that allows decision makers to develop, modify and control decision models. Finally, there is the user interface that connects the system to its operator. It is important to emphasise that the users are the decision-makers, and the system is merely a support tool to ensure that their decisions are made with as much information as possible. In the same vein, one of the strengths of the system is its ability to identify interdependencies between data and, consequently, to anticipate cascading effects, which would allow for significant improvements in the planning and preparedness process for disasters and emergencies (ibid., pp. 28-31).

Finally, there are risk monitoring systems. As mentioned above, Smart Cities live on the data they receive from the massive deployment of IoT devices and sensors that measure everything from traffic flow to temperature, air quality or weather conditions. Risk monitoring systems harness this data and, using advanced data analytics and ML, are able to identify patterns and anomalies that could indicate potential risks. In addition, this technology is sometimes integrated with CCTV cameras and facial recognition systems, providing even more information (ibid, pp 31-32). However, while information is everything in a digitalised and networked society, the price to be paid for it can never be the human rights of individuals, some of which have been highlighted in section 2.1 above. Consequently, the potential risks and vulnerabilities of the technological systems will be analysed in the following section.

## 3.1 LEGAL, ETHICAL AND SOCIAL ISSUES

### 3.1.1. SMART CITY DIGITAL TWIN

Smart cities are being implemented to alleviate the various problems of modern cities and improve the quality of life of their citizens. The sectors that are increasingly implementing smart devices are diverse, ranging from health and transport to education, governance, and the economy. However, as mentioned in previous sections, there is a broad consensus that smart cities are being designed and implemented without the necessary security guarantees (Asif et al., 2022, p. 3; Alamer & Almaiah, 2021, p. 722; Al-Turjman et al., 2019, p. 8). Crime and urban spaces have always been linked, and the emergence of smart cities is no exception (Kitchin & Dodge, 2019, p. 48). Certainly, digital transformation brings with it risks that can have serious implications for cities (Pandey et al., 2019, p. 2). In addition, the proliferation of electronic and smart devices provides attackers with infinite entry points through which they can compromise a city's systems (ibid, p. 5). In general, cybercriminals seek to exploit four vulnerabilities: (1) weak data encryption and software security, (2) use of old and poorly maintained systems, (3) many interdependencies and large and complex attack platforms, and (4) human error, which is often the weakest and most easily exploited step (Kitchin & Dodge, 2019, pp. 49-51).

The infrastructure and assets that can be targeted in a smart city are many, but Toh (2020, pp. 98-101) narrows the list down to the following:

- **Water distribution:** Water is a very important asset to the life and economy of a city. Cybercriminals could gain access to and control of the systems that manage water distribution, with very serious consequences for society.
- **Power grid:** Many of the activities and applications of a smart city are powered by electricity, so it is important to ensure that its production is not interrupted and, if it is, that there are sufficient reserves to meet demand until normal conditions can be restored.
- **Connectivity:** Smart City devices are interconnected, making their connectivity one of the most important assets to protect in a city. The infection or intrusion of a single device could have a cascading effect, bringing down other devices or even the entire system, bringing the city to a complete standstill.
- **Data:** Most smart city applications store large amounts of data. Although not always personal data, they are likely to store and process this type of data about their citizens, so it must also be protected from intrusion through access controls, quality authentication and strong encryption.
- **Financial assets:** Cities are often the economic engine of a country, and one of the goals of smart cities is to improve the economy. Therefore, protecting a city's financial assets is a high priority, whether they are the personal assets of citizens, businesses, or the government.
- **Services:** Services are of great value to a city and its citizens. Indeed, one of the goals of smart cities is to improve the quality of these services. However, services are also often targeted because of the disruptive impact they can have.

Having described the potential targets of attack, it is also worth looking at the different types of attack that could occur in the context of a smart city. In this regard, Al-Turjman et al. (2019, pp. 6-9) summarise the main security threats as follows:

- **Denial of Service:** the aim of this attack is to interrupt the communications of a given system by sending excessive requests to a server until its processing capacity is saturated, causing it to stop working and deny the services offered.
- **Malware:** this is malicious software that exploits vulnerabilities in devices such as cameras to control them and sometimes gain access to confidential information.
- **Eavesdropping:** this is an attack aimed at listening to the communications of third parties without their consent, thus violating the confidentiality of communications and causing an invasion of privacy that could lead to personal and/or economic damage.
- **Masquerading:** stealing information by masquerading as an apparently legitimate and trustworthy device.
- **False Information**: attackers attempt to send false information to devices in order to disable the system. The fact that many decisions are made on the basis of information provided by devices makes this attack particularly dangerous, with serious consequences for a city and its inhabitants.
- **Botnets:** These are infected networks that attempt to modify and/or destroy information on devices such as routers, webcams, IP cameras, printers, etc. From this infected network, DDoS (Distributed Denial of Service) attacks could be launched, which could also bring a city to a standstill.

Due to the wide range of risks and vulnerabilities that exist in a smart city environment, security must be designed and implemented holistically (Toh, 2020, p. 97), protecting devices, systems, and networks at the same time (Alamer & Almaiah, 2021, p. 722) to provide security and trust to a city's most valuable asset: its people. However, traditional security measures such as access control, two-factor authentication, firewalls, antivirus, encryption, etc., while necessary, are not sufficient and will eventually become obsolete. Instead, a systemic approach must be implemented through security-by-design (Ahmad et al., 2022, p. 21; Kitchin & Dodge, 2019, p. 57; Pendey et al., 2019, p. 9). The aim of this methodology is to make security a fundamental and essential part of the design of smart devices and smart cities, with all products being tested before they go on sale. Similarly, apart from investing in security solutions such as cryptography or blockchain technology, emphasis should also be placed on training and raising awareness of people in the field of security.

Security is not the only issue that arises in a smart city environment. Data storage is also susceptible to risks and vulnerabilities that affect privacy (Pandey et al., 2019, p. 7; Cui et al., 2018, p. 46138), a fundamental right that is considered a quintessential right in the digital age. Most smart city applications will capture, collect, process, and analyse a huge amount of data to generate knowledge that will be used to create higher quality services. In our case, PANTHEON will be used to plan and train for disaster scenarios to improve response and enable more informed decision making. However, this process may involve the processing of personal data, which logically must be protected (Toh, 2020, p. 99). Information such as the health status, identity, location, and lifestyle of citizens are examples of information that may put people's privacy at risk (Al-Turjman et al., 2019, p. 8).

Kitchin (2016, p. 4) refers to the datafication of the city and society as an expansion in the amount and scope of data generated about people and places. Indeed, he criticises this approach to urban science, arguing that a city is a complex, multifaceted and contingent set of relational systems, full of complex problems that are not easily measurable. He also identifies five reasons why smart cities could be detrimental to privacy (Kitchin, 2016, pp. 5-9):

- **Datafication, dataveillance and geosurveillance:** The ubiquity of digital transactions with the inevitable use of identifiers containing Personally Identifiable Information (PII) makes it impossible not to leave a digital footprint. Geolocation and surveillance by CCTV cameras could lead to serious violations of many of the rights mentioned in section 2.1 above.
- **Predictive inference and harm:** Predictive modelling from big data, even if it does not collect PII, is capable of discovering characteristics of an individual that are considered to be particularly sensitive personal data, such as sexual orientation, political inclinations or religious beliefs.
- **Anonymisation and re-identification:** The development of new computational techniques can facilitate the re-identification of apparently anonymised or aggregated databases. For example, reverse-engineering strategies could make it possible to identify an individual by linking a pseudonym to another type of data.
- **Reduced control:** The sheer volume of data collected makes it virtually impossible to maintain control over it, i.e. to know what personal data is being collected, with whom it is being shared, how it is being processed and for what purposes.
- **Notice and consent:** Notice and consent are the cornerstones of data protection and the right to privacy. However, the scale of data collection makes the right to notice and consent infeasible.

*Table 3: Main privacy breaches and harms*

| Domain | Privacy Breach | Description |
|---|---|---|
| **Information collection** | Surveillance | Watching, listening to, or recording of an individual's activities |
| | Interrogation | Various forms of questioning or probing for information |
| **Information processing** | Aggregation | The combination of various pieces of data about a person |
| | Identification | Linking information to individuals |
| | Low level security | Carelessness in protecting stored information from leaks and improper access |
| | Secondary use | Use of information collected for one purpose for a different purpose without the data subject's consent |
| | Exclusion | Failure to allow the data subject to know about the data that others have about her and participate in it handling and use, including being barred from being able to access and correct errors in that data |
| **Information dissemination** | Breach of confidentiality | Breaking a promise to keep a person's information confidential |
| | Disclosure | Revelation of information about a person that impacts the way that other judge her character |
| | Exposure | Revealing another's nudity, grief, or bodily functions |
| | Increased accessibility | Amplifying the accessibility of information |
| | Blackmail | Threat to disclosure personal information |
| | Appropriation | The use of the data subject's identity to serve the aims and interests of another |
| | Distortion | Dissemination of false or misleading information about individuals |
| **Invasion** | Intrusion | Invasive acts that disturb one's tranquility or solitude |
| | Decisional interference | Incursion into the data subject's decisions regarding her private affairs. |

*Source: Kitchin, 2016, p. 6*

We are therefore in a situation where massive amounts of data are being collected in multiple contexts, sometimes without the consent of the data subject. In addition, the use of predictive modelling poses a risk of re-identification of anonymised data. Consequently, if ethics are not proactively considered in projects involving data, this could exaggerate vulnerabilities and ultimately cause serious harm to individuals. For this reason, people developing such projects should adopt a privacy-by-design approach, that is, integrate privacy protection throughout the lifecycle of systems that collect data (Hiller & Blanke, 2017, pp. 334-336). One way to do this is to work on a privacy impact assessment (PIA), which aims to address risks that may affect data subjects. This should include (1) the categories of data to be collected, (2) the type of processing to which the data will be subjected, (3) an identification, assessment and prioritisation of potential risks and threats, (4) the mitigation measures to be taken, and (5) constant monitoring throughout the lifecycle (Martín & Kung, 2018, p. 109). In this sense, the use of the LINDDUN framework, based on the Data Flow Diagram (DFD), is highly recommended as a threat analysis methodology in the context of privacy by design, with the aim of achieving privacy-preserving systems (Robles-González et al., 2020, p. 18).

In addition to security and privacy, smart cities also face several social challenges. The hyper-connected and digitised world that is being built typically rewards high levels of digital connectivity and penalises its absence, resulting in a digital divide that negatively impacts disadvantaged groups (Calvo, 2020, p. 144). These disadvantaged groups primarily include the elderly and those with low levels of education (Chang, 2021). Indeed, digital solutions require a high level of digital literacy that not all people have, and some cannot even afford. Failure to take this into account could have a negative impact on democracy and increase inequalities and discrimination in urban areas (Ziosi et al., 2022). A related aspect is also racist, misogynist, and homophobic biases, to name a few, that reproduce certain algorithms and the impact this has on equality and discrimination (Calvo, 2020, p. 144). Finally, another important issue is social control, as smart cities have increased its level, and this raises important ethical and legal dilemma. There is certainly a risk that smart cities will become tools of surveillance and social control, violating human rights in the name of security (Ziosi et al., 2022).

Given the potential issues that could arise in the context of a smart city, it stands to reason that a digital twin based on it would be a replica that inherently replicates them. No matter how much data we have at our disposal, it is unlikely that a digital twin can be created that is identical to reality, as there are variables such as social capital or culture that cannot be measured. There is also a risk of giving too much authority to this technology when its representations are not entirely accurate, are biased or manipulated, and could even be hacked. For this reason, it is recommended that its use should only be of an assistive nature, with individuals being the ultimate decision-makers (Helbing & Argota Sánchez-Vaquerizo, 2023, pp. 81-85). Issues such as cyber threats, false positives or negatives, misleading patterns, calibration and/or validation problems, or misinterpretation are common in digital twins. Similarly, risks of privacy and human rights violations, data misuse, discrimination and inappropriate simplification should also be of concern and would need to be addressed in the design, development, and implementation of the technology (Helbing et al., 2021, pp. 3-4). Consequently, misuse of this technological approach could cause serious harm to individuals and society, so its potential risks need to be carefully identified, investigated, and assessed.

### 3.1.2. INTERNET OF THING SYSTEMS

IoT systems deserve special mention because they are the foundation on which smart cities are built and, like the SCDT, are a fundamental component of the PANTHEON platform. Similarly, security and privacy cannot be neglected in the context of a massive deployment of IoT devices. These devices will collect a lot of data and this data will be shared within the system. Having so many interconnected devices transmitting and sharing data will inevitably have ethical and socio-legal consequences (Righetti et al., 2018, p. 391). While IoT technologies can solve many problems, they also raise the following ethical and legal concerns (Tzafestas, 2018, p. 106):

- Privacy
- Data security
- Usability of data
- Data user experience

- Trust
- Safety

In terms of security, the implementation of IoT systems is a complex and challenging task, as there are many vulnerabilities and the expertise required to produce and secure the system is very diverse, requiring the collaboration of a multidisciplinary approach (Wheeler et al., 2020, p. 66). In addition, each layer of architecture is exposed to risks that may be common or specific to each layer (Mahmoud et al., 2015).

For example, the device layer is exposed to interference by its wireless signals. Similarly, there is a risk of sensors being intercepted by attackers who can take control of them. Finally, one of their greatest vulnerabilities is their low memory and processing power, which makes them difficult to defend against (see Table 4 for more details on device layer threats). Part of the solution to the above lies in the use of strong encryption and authentication.

*Table 4: Attack surface of IoT system edge devices*

| Attack surface | Description | Breach Consequence |
|---|---|---|
| **Network communication protocols and interfaces** | Devices such as sensors and actuators need to be able to send information to and receive commands from a central system. This interaction is achieved using wireless networks such as WiFi or Bluetooth. | An attacker who succeeds in intercepting network or device traffic will have control over it, with the ability to poison the system. This would compromise the integrity of the analysis and the decisions based on it. |
| Operating system | Devices require an operating system to manage the software and network that enables them to function. | Security bugs in the operating system could compromise the entire device, as well as access to its network traffic. |
| **Update mechanism** | One of the best practices is to enable upgrade mechanisms for deployed systems and components. | The device could be compromised if an attacker is able to deliver a malicious update to the system that has not been properly validated. |
| **Network accessible software or services** | This includes the administrative web applications that run the device. | Compromises data, device management or user authentication credentials, and the device. |

*Source: Wheeler et al., 2020, p. 24.*

The gateway layer is vulnerable to various types of attack, including DoS or DDoS attacks and man-in-the-middle attacks (see Table 5 for details of gateway layer threats). To mitigate these risks, good communication protocols must be established, and software must be implemented to respond to situations and system behaviour that are considered abnormal. Finally, if someone with malicious intent gains access to the cloud layer, they could gain access not only to this layer, but also to the previous two layers (see Table 6 for details of threats at the cloud layer) (Mahmoud et al., 2015).

*Table 5: Attack surface of an IoT Gateway*

| Attack surface | Description | Breach consequence |
|---|---|---|
| Network communication protocols and interfaces | Gateways use wireless network communication protocols such as WiFi or Bluetooth to communicate with sensors and actuators. The TCP/IP protocol is typically used to communicate with the back-end cloud. | In this case, there are two attack scenarios: gateway communication with sensors/actuators and communication with the cloud via the Internet. In both cases, the compromise of network traffic could affect the integrity and confidentiality of the entire system. |
| Messaging protocols | These protocols are used to define and manage the transfer of messages and data between the gateway and the devices. | Unauthorised access or modification of data and credentials used by the system. |
| **Operating System** | Same as Table 4. | Same as Table 4. |
| **Update mechanism** | Same as Table 4. | Same as Table 4. |
| **Administrator/management interfaces or applications** | These are applications or management interfaces exposed to external access by the device. | Compromise of data, device administrator or user authentication credentials, and the device. |
| **Basic input/output System (BIOS)** | They are used to perform hardware initialisation during the boot process. | Physical access to the gateway is required to carry out an attack, but if successful it could give the attacker full control of the gateway. |

*Source: Wheeler et al., 2020, p. 24.*

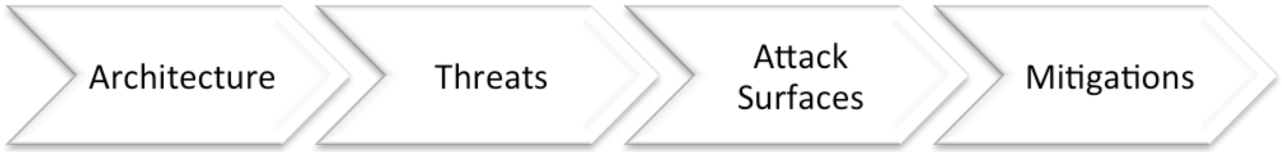
*Table 6: Attack surface of the Cloud*

| Attack surface | Description | Breach consequence |
|---|---|---|
| **Network communication protocols and interfaces** | The cloud is based entirely on the concept of software applications and services delivered over the Internet. Network connectivity is essential and must always be connected, which means it can be attacked by malicious traffic at any time. | For hosted applications on shared or private cloud infrastructure, compromise of host systems, servers, data, or authentication credentials could lead to malicious control of the entire IoT ecosystem for that solution or service, including cloud applications and infrastructures, gateway, and devices. |
| **Network infrastructure** | It includes the network equipment used to connect, | Same as above. |

| | | |
|---|---|---|
| | protect, and access the various machines that make up the cloud architecture, such as firewalls, bastion hosts, switches, routers, etc. | |
| **Messaging protocols** | Same as Table 5. | Same as above. |
| **Operating system** | Same as Table 5. | Same as above. |
| **Update mechanism** | Same as Table 5. | Same as above. |
| **Administrator/management interfaces or applications** | Same as Table 5. | Same as above. |
| **BIOS** | Same as Table 5. | Physical access to the physical server that hosts the virtualized servers that cloud computing application use. An attack here is not as easy to perpetrate or attempt as those previously described, but it's still a concern. |

*Source: Wheeler et al., 2020, p. 25.*

Having seen some of the range of threats that can affect an IoT system, it is useful to plan its security (see Section 5 for more details). One of the most popular methods for doing this is ATASM: (1) Architecture, (2) Threats, (3) Attack Surfaces and (4) Mitigations (see Figure 5).



*Source: Schoenfield, 2015, p. 32.*

Figure 5 – ATASM security process

In the same vein, Wheeler et al. (2020, pp. 71-82) propose the following process for designing the security of an IoT system:

1. **System architecture analysis:** the analysis should include a conceptual view of the elements of the system and their interconnections, their decomposition into logical functions and layers, including the flow of data between modules, and finally the data and features of most value to the system.
2. **Threat analysis:** this is a list of potential vulnerabilities or points of attack in the system. The analysis should include (a) the assets of the system and how they can be exploited, (b) the types and methods of attack to which the system is susceptible, (c) the types of adversaries, their motives, and capabilities, and (d) the interfaces exposed to vulnerabilities and points of attack.
3. **Threat classification***:* Once threats have been identified, they can be categorised as follows
   a) Accepted threat: accepts the risk and does nothing to mitigate it.
   b) Transferred threat: document and transfer the risk to third parties such as the user.

c) Mitigated threat: add features to the system to prevent the attack.
d) Detected threat: add features to the system that react to the attack but do not prevent it. For example, notification that a security breach has occurred.

4. **Incorporate corrective or mitigating measures**: this involves designing countermeasures and controls into the system that manage and mitigate the identified risks.
5. **Repeat the process:** each time the architecture is changed, the process must be restarted to identify new risks and threats that may arise as a result of the change.
6. **External system security review:** this involves bringing in external architecture security specialists (who may be from the same consortium) to review the design for vulnerabilities that have been overlooked by the lead designers.

Indeed, the best way to achieve maximum system security and protect users' rights is to adopt a security-by-design approach, i.e. to implement security measures in the system's software and hardware throughout its lifecycle, from the design phase to the deployment phase (Atlam & Wills, 2020, p. 130). To this end, Mahmoud et al. (2015) cite several principles that should guide the process of designing, developing, and deploying IoT systems:

- **Confidentiality:** data must be secure and available only to authorised users.
- **Integrity**: data must be guaranteed to be accurate, to come from the true sender and not to have been tampered with.
- **Availability**: data must be fully available to authorised users.
- **Authentication**: each object in the IoT system must be properly identified and authenticated with respect to other devices. A mechanism must be implemented to authenticate devices for each communication.
- **Heterogeneity**: communication protocols must be designed to work on any device and in different situations. Adequate key management and security protocols must be ensured.
- **Key management system**: communication and data exchange between devices must be encrypted to ensure data confidentiality.
- **Security policy:** a clear security policy must be established.
- **Awareness**: humans are the weakest link in a system, so users should be informed and trained about system risks and the corrective actions they can take to prevent them from materialising (Righetti et al., 2018, p. 391). For example, system providers should encourage users to change default passwords to ones that meet appropriate security requirements (Altam & Wills, 2020, p. 131).

To the above principles, we could also add some of those mentioned by Altam & Wills (2020, pp. 131-132):

- **Tamper resistance**: physical access to IoT devices should be prevented to prevent attacks.
- **Updating**: security patches and updates should be implemented to mitigate potential new risks or vulnerabilities identified during the ongoing process of system monitoring.
- **Pentesting**: Penetration testing should be performed regularly on both software and hardware to test the system and identify and mitigate vulnerabilities.
- **Resilient design**: Although practically impossible, efforts should be made to design a system that is resilient to unforeseen situations where there is no connectivity.

While it is true that the security and privacy of a system are inextricably linked, this does not preclude a specific mention of privacy and personal data protection. Some of the threats to privacy posed by IoT systems include the following (Atlam & Wills, 2020, pp. 138-140):

- **Identification**: IoT devices sense and collect various types of data about their users and their interactions with the environment. Identification is a threat that involves linking an identifier to private information about the individual.
- **Location and tracking**: Specifying and recording a person's location through technologies such as GPS or internet traffic.
- **Profiling**: collecting and processing data about an individual's activities and actions over time and classifying them according to certain characteristics.
- **Inventory raiding**: the unlawful collection of information about the existence and characteristics of things found in a particular place.
- **Linking**: the disclosure of personal information by combining separate data sources.

When we talk about a personal data breach, we refer to (a) *destruction*, where the data no longer exists; (b) *loss*, where the data still exists but the controller has lost control of or access to it; (c) *damage* by alteration or corruption; and (d) *unauthorised or unlawful processing*, which consists of both disclosure and access to personal data by unauthorised recipients. Sharma (2020, pp. 95-96) categorises personal data breaches in two ways. One is the CIA triad (Confidentiality, Integrity, Availability):

- **Confidentiality breach**: unauthorised or accidental disclosure or access to data.
- **Integrity breach**: accidental or unauthorised modification of personal data.
- **Availability breach**: accidental or unauthorised destruction of or loss of access to personal data.

On the other hand, we could also classify them according to their mode of execution:

- **Physical**: involves the physical destruction or theft of electronic systems such as computers or external hard drives.
- **Electronic**: where a cybercriminal gains unauthorised access to and/or control of, or disruption to, the processing of data. This can be done through malware, phishing, password (or brute force) attacks, ransomware, DoS, or DDoS attacks, and coordinated attacks involving multiple types of the above.

The damage that can be caused by privacy breaches is numerous and varied, including loss of control over one's own data, limitation of user rights, discrimination, identity theft or fraud, economic losses, unauthorised reversal of anonymisation, damage to reputation, loss of confidentiality of personal data protected by professional secrecy, and other economic and/or social disadvantages for the individual (Sharma, 2020, pp. 96-97). The formula for mitigating, minimising, and preventing the above harms lies precisely in a privacy-by-design approach, in which privacy is fully integrated into the components that make up the system (Atlam & Wills, 2020, pp. 140-141). In this sense, designers, and architects of IoT systems must think carefully about the privacy implications of their systems and design them with appropriate privacy controls that guarantee users' rights (Wheeler et al., 2020, p. 235). Some of these controls include (Atlam & Wills, 2020, pp. 140-141; Wheeler et al., 2020, pp. 235-240):

1. **Data minimisation**: the aim is to reduce the collection of personal data to that which is strictly related to the service to be provided. The same applies to data retention and the retention of only the personal data necessary to provide the service.
2. **Access controls**: There are various access control measures that can be applied to ensure privacy. We highlight the following:
    - Restrictive permissions: files containing personal and/or sensitive data should be restricted to authorised personnel only.

- Default login to systems and applications: all accounts on a system should require login or other verification and authentication systems.
- Use of strong passwords for all accounts: Account passwords should be strong and changed frequently. The system should include a mechanism to ensure that this is the case, thereby preventing the user from making security mistakes that could compromise the system. Passwords should be required to be at least 8 characters long, although 12 characters is recommended. Passwords should contain a variety of character types, including upper and lower case letters, numbers and special characters. Users should be made aware of the risks of using weak passwords and, where possible, the use of passwords should be audited.
- Two-factor authentication: Accounts with privileges, such as access to a significant amount of data, should have two-factor authentication to ensure greater protection.
- Review and delete accounts: An annual review of accounts and privileges is recommended to remove old accounts that are no longer in use or to remove privileges for people who no longer need access to certain files.

3. **Anonymisation**: This is the removal of PII from the dataset to protect the identity and privacy of individuals. There are several ways in which data can be anonymised, including the following:
    - Deletion: This consists of directly deleting the PII from the record. It should be noted that although this is very effective, it may reduce the value of the record.
    - Masking: This consists of changing the characters of the fields that contain PII to characters such as the asterisk or space.
    - Generalisation: This consists of changing fields containing PII to general terms that do not reveal personal information. For example, GPS coordinates can be changed to the name of the city.
    - Tokenisation: Consists of changing fields containing PII to a symbolic value. This can be useful if you want to give access to or share records with a partner or affiliate. It is possible to have a master file that contains the relationship of the symbolic codes to the actual PII. Of course, this master file will need to be protected by additional technical and organisational measures.
4. **Encryption**: Appropriate cryptographic techniques should be used to ensure the privacy and confidentiality of data and communications. There are various techniques available, and the choice should be made according to the risks identified and the human and material resources available. Some of the existing techniques are homomorphic encryption, which is a special encryption algorithm that allows some computational operations to be performed without having to decrypt the data. One can also resort to differential privacy, which is a new technique that obfuscates a data set by introducing a calculated number of errors, allowing the requesters to make the relevant statistical inferences.
5. **Data retention and deletion policies**: the people designing the system must be aware of the information they are collecting, voluntarily or involuntarily, or allow the system to collect and manage that information with due care. This means implementing policies that define what data is collected, what data is retained, and for how long. Ideally, all data elements, including those that are not considered PII, should be reviewed in terms of their value to the system. The value they bring to the system and the privacy risks they pose should also be assessed.
6. **Awareness**: One of the biggest privacy problems is that there is still a lack of public awareness of the risks and threats posed by certain technologies. Awareness should be raised of these, as well as of the self-protective measures we can take to avoid harm.

### 3.1.3. ARTIFICIAL INTELLIGENCE

As described in Section 2.3 and in Deliverable D10.3, AI brings many benefits to society, but it also poses significant challenges that need to be addressed. The harms that AI can cause can be both tangible and intangible. This can have consequences for the safety and health of individuals, including the loss of human life and damage to property. It can also have consequences for privacy and data protection, the restriction of rights such as freedom of expression or assembly, and serious implications for dignity and the non-discrimination regime. Undoubtedly, the use of technologies such as AI can undermine the fundamental values on which the EU was founded and have serious ethical, legal, and social implications (European Commission, 2020, pp. 10-11).

It cannot be disputed that AI has the potential to speed up and improve the decision-making process by analysing vast amounts of data collected by sensors and other devices in smart cities (Yigitcanlar et al., 2020, p. 10). Focusing on the objectives of the PANTHEON project, AI - and more specifically artificial neural networks - was used together with satellite imagery after an earthquake in the city of Palu (Indonesia) in 2018 (Syifa et al., 2019, p. 15). These networks, with a computational method for data classification, are able to acquire, present and compute data for mapping with the aim of making predictions (ibid, p. 6). Furthermore, AI-based systems can detect potential terrorist threats by monitoring communications (Yigitcanlar et al., 2020, p. 13), which would allow for better prevention of violent radicalisation.

However, AI poses risks such as erratic and even unethical decision-making, as well as threats to rights such as privacy, human dignity, equality, or security. It can also exacerbate inequalities that already exist in society, particularly in cities, and threaten certain human jobs (Meek et al., 2016, p. 689). For example, recent research has shown that some facial recognition software is inherently racially biased (Yigitcanlar, 2020, p. 5). It is therefore logical to conclude that many of the projects involving AI will have some ethical implications. The problem is that, in most cases, this impact will be discovered either during the use of technology or during its development. In contrast, an ethics by design approach aims to address AI challenges before they can materialise (Brey & Dainow, 2020, p. 7). It is precisely this approach that the PANTHEON consortium intends to take in relation to the system it will design, develop, and implement.

One of the most comprehensive pieces of work done on the design of ethical AI systems is the guidance produced by the European Commission's High-Level Group on Artificial Intelligence (see Deliverable 10.3 for more information). According to the Group, for an AI system to be trustworthy, it must fulfil three fundamental components. First, it must be legal, i.e. it must comply with the relevant existing legislation. This includes EU primary law (the EU Treaty, the EU Treaty on the Functioning of the EU, and the EU Charter of Fundamental Rights), its secondary law (the GDPR, the Product Liability Directive, anti-discrimination directives, consumer protection directives, etc.), UN and Council of Europe human rights treaties, as well as relevant national legislation. On the other hand, the AI system must be ethical, as legislation often comes late, and it is best to ensure that AI system providers voluntarily adhere to ethical codes of conduct. Finally, the AI system must be robust and guarantee that its operation does not cause unintended harm (HLEG-AI, 2019, pp. 6-7).

In addition to the fundamental elements mentioned above, the HLEG-AI has also identified four ethical principles to which all AI systems to be developed in the EU must adhere. These are: (a) respect for human autonomy, (b) prevention of harm, (c) justice, and (d) explicability. It is possible that some conflicts may arise between these principles, creating dilemmas for which there is currently no solution. It is up to designers and developers to consider them and make relevant decisions, which must necessarily be based on clear evidence and justified and proportionate use cases (HLEG-AI, 2019, pp. 12-13). Finally, these four ethical principles lead to the seven requirements for an AI system to be considered trustworthy. These are as follows:

1. **Human agency and oversight**: AI systems should only support decisions made by humans. No decision should be made on the basis of automated processing alone. Human oversight of systems is a good way to ensure that the system does not undermine human autonomy or cause adverse effects.
2. **Technical robustness and safety**: AI systems must be sufficiently protected, both in terms of software and hardware, against threats and vulnerabilities that can be exploited by cyber criminals. Appropriate measures must be taken to prevent these risks. The data used to train the AI must be accurate, otherwise the system could cause harm to people. A contingency plan should be put in place for cases where adverse effects are likely to occur.
3. **Privacy and governance**: AI systems must ensure privacy and data protection throughout their lifecycle. The data on which AI is trained may have biases that lead to discrimination, so it is critical that the data is of the highest possible quality and integrity. Protocols need to be established on who can access the data and under what conditions. Measures such as anonymisation, pseudo-anonymisation, encryption and aggregation must be taken to protect users' rights. Data must be collected, stored, and used in a way that allows human auditability.
4. **Transparency**: The data sets and decision processes of an AI system must be recorded and documented to ensure traceability. Users operating the AI system must know at all times that they are interacting with an AI and understand the decisions it is making. The objectives, limitations, benefits and risks of the AI system and its decisions should be openly communicated to users, including instructions on how to use the system appropriately.
5. **Diversity, non-discrimination, and fairness**: Training an AI system with biased data can exacerbate discrimination through unfair distortions. These biases should be minimised, mitigated and, where possible, eliminated from the system. A good way to do this is to involve a culturally and academically diverse team in the design, development, and implementation of a system. The design of the system should be user-centered, so that everyone can use it, regardless of age, gender, or ability. It is advisable to involve the different stakeholders who may be affected by the system in the design and development process of an AI system.
6. **Social and environmental well-being**: the system must be sustainable and environmentally friendly. Resource use and energy consumption should be examined to adapt to the least environmentally damaging option. While it is true that AI aims to enhance human capabilities, it is also possible that it will degrade them, so one must be aware of the possible social consequences of the system being designed. Similarly, its impact on institutions, democracy, and society in general needs to be considered.
7. **Accountability**: it should be possible to audit the system, i.e. to evaluate the algorithms, the data, and the design process. The audit should be both internal and external to the organisation/consortium. Potential negative impacts should be identified, evaluated, documented, and minimised. In the case of conflicting requirements, the decision maker should make decisions based on evidence and appropriate justification, considering the risks and benefits of the system.

### 3.1.4. UNMANNED AERIAL VEHICLES

Drones in Athens metropolitan are used by the Emergency Agencies (e.g., Hellenic Police, Fire Service) in the following ways:

- Security and surveillance of sporting events (i.e., Authentic Athens Marathon, football matches, other major sporting events etc.).
- Live image transmission to the Police Operations Center of large events, police or fire operations, forest/urban fires and protests/public assemblies/gatherings.
- Search and rescue of missing persons.

In order to legally and safely conduct flights and mitigate risks, reckless and malicious drone use, current legal and SOPs require:

- Issuance of pilot certification according to EASA (European Aviation Safety Agency).
- Drawing up a flight plan before each flight.
- The macroscopic control of the required equipment.
- In the case of flights during sporting events and gatherings, where live image transmission is requested/required, the issuance of a decision by the Joint Coordination Center of Operations & Crisis Management of the Hellenic Police Headquarters and the respective Primary Prosecutor, of the area of implementation is primarily required of flight, regarding the installation, processing and operation of a portable surveillance system, according to P.D. 75/2020.
- The submission/preparation of a flight Risk Assessment in the Special Category (Specific Operations Risk Assessment), for each operation/flight.
- Insurance in which a company or the government undertakes to provide a guarantee of compensation for specific loss, damage, illness or death caused by a drone in return for the payment of a specific premium.

It should be noted that drones in Greece, as in any European city, are subject to the following regulations:

- 945/2019 Regulation of the European Commission.
- 947/2019 Regulation of the European Commission.
- 639/2020 Regulation of the European Commission.
- Additional rules/prohibitions established by the national authority (Civil Aviation Authority - Civil Aviation Service).

# 4. PANTHEON'S SECURITY AND PRIVACY BY DESIGN FRAMEWORK

## 4.1. INTRODUCTION

PANTHEON project addresses a wide area where the rights of persons are to be considered. By the usage of advanced technologies, including AI, the project should address security of data, privacy protection, ethics, social acceptance, and legal aspects.

The above-mentioned aspects should be addressed in early stages of the project implementation and the focus should be on strategy and security measures, including security and privacy by design; assessment of the technology usage, risk management, administrative measures and compliance with standards. These measures will be further explained.

### 4.1.1. ETHICAL ARGUMENTS

- The universality of principles and values: A technology is developed based on values or principles that are assumed to be universal.
- Technological determinism: Technologies have both positive and negative effects. The determinism relates to the a priori knowledge about the balance between positive and negative effects.
- Neutrality of technology: The technology is not designed to be good or bad. It is neutral. The particular usage of a technology can obtain positive or negative effects.
- Precedence: New technologies are based on precedent ones and inherit the characteristics.
- Change of ethical values: Technologies change our values leading to moral progress or moral decline.

### 4.1.2. THE CIA TRIAD

The purpose of identifying the security criteria is to determine the threats that could affect the target system and aim at protecting the identified assets. They are formulated in terms of Confidentiality, Integrity, and Availability, also known as CIA triad (Ingeno, 2018) which is a guiding model in information security. A comprehensive information security strategy includes policies and security controls that minimize threats to these three crucial components. The CIA triad guides information security in a broad sense and is also useful for managing research products and data.

The first component of the Triad is **confidentiality** of all data assets. It is a property of essential elements making them accessible only to authorized users. Confidentiality involves the efforts of an organization to make sure data is kept secret or private. To accomplish this, access to information must be controlled to prevent the unauthorized sharing of data—whether intentional or accidental. A key component of maintaining confidentiality is making sure that people without proper authorization are prevented from accessing assets important to your business. Conversely, an effective system also ensures that those who need to have access have the necessary privileges. Confidentiality can be considered as a function as protection of algorithms describing the management rules and results whose disclosure to unauthorized third parties could be harmful, non-disclosure of processing or a mechanism of a confidential nature. Regarding information, confidentiality means the protection of data whose disclosure to unauthorised third parties could be harmful, and the non-disclosure of data of a confidential nature.

The second component is **integrity** which is a property defining the accuracy and completeness of the essential elements. Integrity involves making sure your system or data is trustworthy and free from tampering. The integrity of your data is maintained only if the data is authentic, accurate, and reliable.

Integrity can be viewed as a function as assurance that the algorithm is correct, or that automated or non-automated processing is implemented according to the specifications, no incorrect or incomplete results from the function. As for information, integrity refers to guarantee that no operating errors or unauthorized uses have impaired the accuracy and exhaustiveness of the data, no corruption of the information.

The third component is **availability** and business continuity. It is a property of essential elements that allows authorized users to access them at the required time. Even if data is kept confidential and its integrity maintained, it is often useless unless it is available to those in the organization and the customers they serve. This means that systems, networks, and applications must be functioning as they should and when they should. Also, individuals with access to specific information must be able to consume it when they need to, and getting to the data should not take an inordinate amount of time. Availability can be viewed as a function as guaranteed continuity of processing services as well as absence of problems linked to response times in the wide sense. In the case of information, availability refers to a guarantee of the proposed availability for access to data (time-to-access and availability timetable), there is no total loss of information; as long as there is a back-up of the information, it is considered to be available. It is assumed that there is a backup, and availability is assessed in terms of the backup function for this information.

### 4.1.3. CLASSIFICATION LEVELS

Table 7: CIA classification levels presents the classification levels for the 3 criteria of confidentiality, integrity, and availability based on the EBIOS (Expression des Besoins et Identification des Objectifs de Sécurité) methodology (ENISA, 2023).[3]

*Table 7: CIA classification levels*

| **Impact level** | | **Availability** | **Confidentiality** | **Integrity** |
|---|---|---|---|---|
| **0** | No impact | No availability need | Public | No integrity need |
| | | The essential element may be momentarily or definitively unavailable without consequence. | All users have unrestricted access to the essential element. | There is no need to guarantee the integrity of the essential element. |
| **1** | Low | Long term | Limited basic | |
| | | The essential element may be unavailable for more than a week but must not be definitively lost. | The essential element can only be accessed by PANTHEON personnel or identified and involved customers. | |
| **2** | Medium | Medium-term | Limited high | Medium integrity need |
| | | The essential element must be available during the week. | The essential element can only be accessed by PANTHEON personnel with a need-to-know. | There is a need to guarantee the integrity of the essential element (medium importance). |

[3] https://www.enisa.europa.eu/topics/risk-management/current-risk/risk-management-inventory/rm-ra-tools/t_ebios.html

| | | | | |
|---|---|---|---|---|
| **3** | High | Short term | EU classified | |
| | | The essential element must be available during the day. | The essential element is reserved for a limited number of persons on a need-to-know basis. | |
| **4** | Very high | Very short term | | Total integrity |
| | | The essential element must be available in real-time. | | The essential element must be perfectly integrated. |

## 4.2. APPROACH PROPOSED FOR PANTHEON PROJECT

To ensure that the PANTHEON system has optimal ethical and legal guarantees, the following directions of work should be considered:

- Security and privacy by design
- Assessment of the technology usage
- Risk management
- Testing for compliance with relevant standards
- Administrative measures

Some guidelines and recommendations for each of the relevant directions of work are presented below.

### 4.2.1. SECURITY AND PRIVACY BY DESIGN

Security and privacy by design (SPbD) is an approach to developing systems, products, and services with security and privacy considerations built into every stage of the design and development process, i.e. throughout the entire development lifecycle. By privacy we refer to data usage, data protection, anonymization, user consent, and control, preserve third-party privacy.

SPbD is aligned with the United States National Institute of Standards and Technology (NIST, 2023). SPbD is made up of six components: Threat model, Privacy Assessment, Vulnerability Management, Secure Release Process, Code Scan and Penetration Test. SPbD will be used in Task 3.7 and will be reported in D3.7 “Overall Architecture and High-level Functionalities”.

### 4.2.2. ASSESSMENT OF THE TECHNOLOGY USAGE

The assessment of the technology usage is a process conducted by security and ethical experts and will be organized in three phases covering eight steps (Burgess & Kloza, 2021). The first phase is the preparation of the assessment process and consists of three steps, namely screening, scoping, and planning. Step one, screening, will consist of a preliminary description followed by a screening of data protection, ethics and social acceptance and privacy. By social acceptance we refer to the degree to which technology is accepted or tolerated. The second step will focus on scoping and will include benchmarking of personal data protection, ethics, social acceptance, and privacy. It will also include scoping of stakeholders and their consultation techniques; appraisal techniques and other evaluation techniques. Finally, the third step consists of planning and preparation, objectives, acceptability criteria, procedures and timeframes, assessors, continuity of the process and revisions.

The second phase is the assessment and consists of three steps. Step four which is a systematic description of the results. Step five focuses on appraisal of impacts; data protection to address the necessity and proportionality of the processing and with the risk to the rights and freedoms of natural persons; ethical assessment; social acceptance assessment; privacy impact assessment and a legal compliance check. And finally, step six, which will include recommendations.

The third phase is the post evaluations and includes step seven, consultation with supervisory authorities, and step eight, revisiting.

A detailed description of the steps can be found in Burgess & Kloza (2021). The assessors can define templates for assessment, or use existing ones defined in this book.

### 4.2.3. RISK MANAGEMENT AND THREATS

Firstly, the way in which threats are considered is explained, then risk assessment and risk management.

The threats are formalized by identifying their characteristics: the attack methods to which the organization is exposed, the threat agents that may use them, the vulnerabilities exploitable on the system entities, and their level.

An attack method is characterized by the security criteria (availability, integrity, confidentiality) that it can violate as several criteria can be affected at the same time. It is associated with threat agents. A threat agent can be characterized by its type or origin, natural, human, or environmental, and by its cause, accidental or deliberate. The intrinsic vulnerabilities of a system arise from the characteristics of the entities it contains. These vulnerabilities can be exploited to attack the security system.

Attack methods are generally selected among the following themes: physical damage; natural events; loss of essential services; disturbance due to radiation; compromise of information; technical failures; unauthorized actions and compromise of functions.

Vulnerabilities can be characterized by their level, representing the possibility of achieving the attack methods that exploit them. The vulnerabilities levels proposed by us are explained in the next table Table 8: Vulnerability levels

*Table 8: Vulnerability levels*

| Level | Accidental Cause | Deliberate Cause |
|---|---|---|
| 0 | Improbable | Unfeasible |
| 1 | Low probability | Needing very considerable means and/or a very high level of knowledge in the field concerned |
| 2 | Medium probability | Needing some degree of expertise and/or specific equipment |

| | | |
|---|---|---|
| **3** | High probability | Possible using standard means and/or basic knowledge |
| **4** | Certain | Possible for anyone |

Characterization of threat agents can be summarised by a single value for each attack method selected, called the attack potential. If the attack methods constitute real risks for the target system, the level of security measures must be consistent with this attack potential. This value is represented by one of the following values defined by us in the following table (Table 9: Level of threat (Attack potential).

*Table 9: Level of threat (Attack potential)*

| **Level** | **Accidental Cause** | **Deliberate Cause** |
|---|---|---|
| ***1*** | ***Accidental*** | ***Random*** |
| **2** | Limited probability | Limited expertise necessary to attack the system or limited access to resources or information |
| **3** | High level of probability | High level of expertise necessary to attack the system or high level of access to resources or information |

The risk level is determined according to the following formula:

Risk = Impact X Attack Potential x Vulnerability Level

The risk that a particular element is subject to is a combination of the level of the threat (attack potential), the level of vulnerability (opportunity) of that element to the threat, and the level of impact (damage that can be caused) to the element should the threat materialize.

A risk score is calculated for every risk [by multiplying the ‘Impact level’, 'Attack Potential' and ‘Vulnerability Level’ from which a risk level is then determined (shown in tables 4, 5 and 6).

The result of the multiplications of the three elements mentioned, is presented in the next table: Table 10: Risk score matrix

*Table 10: Risk score matrix*

| **Attack potential (threat)** | ***1*** | ***1*** | ***1*** | ***1*** | ***1*** | ***2*** | ***2*** | ***2*** | ***2*** | ***2*** | ***3*** | ***3*** | ***3*** | ***3*** | ***3*** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Opportunity (vulnerability)** | ***0*** | ***1*** | ***2*** | ***3*** | ***4*** | ***0*** | ***1*** | ***2*** | ***3*** | ***4*** | ***0*** | ***1*** | ***2*** | ***3*** | ***4*** |
| **Asset value** | **Risk Score** | | | | | | | | | | | | | | |
| ***1*** | 0 | 1 | 2 | 3 | 4 | 0 | 2 | 4 | 6 | 8 | 0 | 3 | 6 | 9 | 12 |
| ***2*** | 0 | 2 | 4 | 6 | 8 | 0 | 4 | 8 | 12 | 16 | 0 | 6 | 12 | 18 | 24 |
| ***3*** | 0 | 3 | 6 | 9 | 12 | 0 | 6 | 12 | 18 | 24 | 0 | 9 | 18 | 27 | 36 |
| ***4*** | 0 | 4 | 8 | 12 | 16 | 0 | 8 | 16 | 24 | 32 | 0 | 12 | 24 | 36 | 48 |

Based on the risk score matrices, we are creating the levels of risks presented in the next table:
Table 11: Levels of risk

*Table 11: Levels of risk*

| Risk Score | Risk Level |
|---|---|
| **0** | No risk |
| **1 – 4** | Low |
| **6 – 9** | Moderate |
| **12 – 18** | High |
| **24 – 48** | Very High |

### 4.2.4. ADMINISTRATIVE MEASURES

Each technical partner utilizing technologies has identified project-specific risks, centrally monitored, and mitigated. Critical risks have been outlined in the LIST OF CRITICAL RISKS in PANTHEON. Security risks are described in the Security-by-Design chapter of the system architecture (Task 3.7).

Two initial AI-related risks include user scepticism due to legal and privacy concerns and a lack of suitable datasets for AI and ML model training. Mitigation measures involve PIA+ assessments, user-centric approaches in specific tasks, and dedicated ethics workshops. Regarding dataset challenges, technical partners will seek alternatives or collect new datasets in compliance with ethics regulations. A shared risk log is maintained throughout the project by Work Package Leaders and the Project Coordinator, incorporating critical risks and adding new ones, including AI and ethics risks. Additional risks identified after AI questionnaires include potential bias, influencing human decision-making processes, and data privacy concerns. Mitigation measures for bias include rigorous data pre-processing, diverse dataset utilization, testing models before and after deployment, and adherence to trustworthy AI guidelines. Human decision-making influence risk is mitigated by using AI as a complementary tool under human supervision. Privacy risks follow data processing rules based on GDPR, particularly for processing biometric data and transferring data outside the EU.

In addition to the measures already mentioned, the commitment made by all consortium partners to comply with regulations, to follow a Code of Ethics, and to take all the necessary security measures, other measures have been implemented to support the implementation of Project PANTHEON:

- Establishment of a PANTHEON Ethics Committee,
- A Quality Management system that will monitor the quality of both software and written deliverables, including testing and data training,
- A Risk Management System that will include monitoring of AI-related risks,
- A Data Management Plan,
- Ethics protocols followed by PANTHEON partners, as presented in D10.1, D10.2, D10.3 and D10.4.,

- Measures were taken for compliance with rules when dealing with the processing of genetic, biometric, and/or health data.

### 4.2.5. TESTING FOR COMPLIANCE WITH RELEVANT STANDARDS

The standards used to check the system compliance will be:

- ISO/IEC 27000-series (also known as the 'ISMS Family of Standards' or 'ISO27K' for short related to information security)
- ISO 31700-1:2023 (standard on Privacy by Design)
- NIST SP 800-160 Vol. 1 Rev. 1- Engineering Trustworthy Secure Systems

# 5. CONCLUSIONS AND RECOMMENDATIONS

The purpose of this report was to analyse the ethical, legal, and social issues that may arise during the design, development, and implementation of the PANTHEON system/platform, and to make some recommendations in this regard. The integration of human rights in technologies such as the one to be developed by PANTHEON is a necessity and increases the quality of the product. Issues such as security and privacy are often overlooked by users and consumers of technology, who only value its functionality. While it is important for the PANTHEON system to be a useful tool for community-based disaster management, ethical issues should not be an afterthought and should be integrated into the design process so that the product that comes to market is safe and respects the human rights of its users or consumers.

The impact of new technologies on the human rights regime is remarkable. Equality is threatened by issues such as unequal access to technology and the skills to live in a digitalised world. Clearly, the benefits of technological advancement do not reach all groups in a society equally. As a result, systematic discrimination by technologies such as AI is no longer a potential threat, but a reality. Human biases and prejudices have inevitably been transferred to new technologies through insufficiently conscious design. Discrimination by algorithms against already disadvantaged and marginalised groups is an issue that has to be taken seriously in our society and must be minimised as far as possible.

In the same vein, the digitalisation of society has allowed for better and greater surveillance of citizens. While this can bring benefits in areas such as the prevention of terrorism, it also entails risks that we cannot afford. In particular, social control technologies can represent a serious violation of freedom of opinion and expression, as well as freedom of assembly and association. Digitalisation has also multiplied the challenges and threats to security and privacy. This process is a phenomenon that brings a multitude of benefits in different areas, and it would be a mistake to slow it down. However, the creation of new digital spaces must be based on human rights to ensure that they are secure and private. Otherwise, people would be exposed to serious violations of their human rights. While there are already legal frameworks in place to deal with the phenomenon of emerging technologies, there is also a need to create new regulations to address the new challenges posed by disruptive technologies. The EU Artificial Intelligence Act and the European Declaration on Digital Rights and Principles are examples of legislative developments in this regard.

For the time being, the PANTHEON project and its resulting product do not appear to pose serious human rights risks. The PANTHEON system is intended to be a support tool for key actors in the event of disasters or various emergencies. Within all the applications identified, it has been decided that the system will focus on the pre-disaster phase, serving both disaster management planning and training of relevant actors. Both applications will be performed through digital simulations based on a Smart City Digital Twin, which will be able to predict the evolution of the selected disasters for each pilot area.

Indeed, both the Smart City Digital Twin and other technologies that PANTHEON intends to use (IoT systems, decision support systems, risk monitoring systems, artificial intelligence) have certain vulnerabilities and are exposed to the threats described in this deliverable. Although PANTHEON does not pose a serious threat to the human rights of its users, the consortium has adopted an ethics-by-design, security-by-design and privacy-by-design approach to ensure the rights of its users and of society at large. We strongly believe that this will add great value to the PANTHEON system and therefore, based on the work done during Task 3.5, a catalogue of recommendations (see Table 12) has been drawn up for PANTHEON designers and developers to consider in order to create an ethical, secure and human rights compliant technology.

*Table 12: Ethical, legal, and social design recommendations catalogue*

| Title | Description |
|---|---|
| | **Legal issues** |
| **Human Rights** | Consider the impact that the PANTHEON system may have on the human rights highlighted in this report, namely (1) equality, (2) non-discrimination, (3) security, (4) privacy, (5) freedom of thought, (6) freedom of opinion and expression, and (7) freedom of assembly and association. This can be done through a human rights impact assessment. |
| **Consumer Rights** | Ensure the quality of the product to be developed under the PANTHEON project by complying with the relevant specific regulations, i.e. the Consumer Rights Directive (2011/83/EU) and the Product Safety Regulation (EU Regulation 765/2008). |
| **Climate change** | Consider the environmental impact of the PANTHEON system, particularly in terms of energy consumption. It is recommended to choose energy options and consumption habits that are less harmful to the environment. |
| **Artificial Intelligence** | Consider the UNESCO and OECD recommendations on the use of AI. Consider the recently adopted EU law on artificial intelligence and the ethical recommendations of the EU High Level Group on Artificial Intelligence. At PANTHEON use cases national level, consider Greece's Law 4961/2022 and Austria's National Strategy for the Use of AI. |
| | **Safety and security by design** |
| **Security and Accuracy of the Smart City Digital Twin** | Ensure, as far as possible, that the devices and processes used to generate the Smart City Digital Twin are secure and that the data collected is accurate and confidential. |
| **Security at the Device Layer of the IoT System** | Consider and assess the attacks highlighted in this report (Table 4) but identify potential alternative risks and vulnerabilities that have not been foreseen in this report. Risk analysis should be a constant throughout the PANTHEON lifecycle and mitigations should be implemented accordingly. |
| **IoT System Gateway Layer Security** | Consider and evaluate the attacks highlighted in this report (Table 5), while identifying potential alternative risks and vulnerabilities not foreseen in this report. Risk analysis should be a constant throughout the PANTHEON lifecycle and mitigations should be implemented accordingly. Particular attention should be paid to DoS/DDoS and man-in-the-middle attacks. |
| **Security at the cloud layer of the IoT system** | Consider and assess the attacks highlighted in this report (Table 6), while identifying potential alternative risks and vulnerabilities not foreseen in this report. Risk analysis should be a constant throughout the PANTHEON lifecycle and mitigations should be implemented accordingly. |
| **Security Planning** | The architecture of the PANTHEON system must be studied to identify its most vulnerable features or elements. With this in mind, both possible attacks and the different attack surfaces that a cybercriminal might target should be identified, studied and evaluated. Threats should be classified according to whether they are accepted, transferred, or mitigated. Finally, mitigation measures for the identified risks should be defined. These measures should be addressed during the design of the system, following the principle of security by design. |
| **External security review** | It is recommended that security experts who have not been involved in the security design of the architecture can review the system for potential weaknesses/vulnerabilities that may have been overlooked by the main designers. |

| | |
|---|---|
| | These experts may be from an organisation that is part of the PANTHEON consortium, although it is strongly recommended that experts from outside the consortium also participate. |
| **Device authentication** | Devices in the IoT system must be properly identified and authenticated with respect to the other devices in the IoT system. |
| **Key management system** | Communications and data transfers within the system must be properly encrypted. |
| **Ensure the CIA triad** | Guarantee Confidentiality through data security, Integrity through data accuracy and the establishment of measures to prevent tampering, and Availability of data only to authorised persons. |
| **User awareness** | System users should be aware of both the risks and threats to the security of the system and the corrective and self-protection measures they can take. |
| **Updating the system** | The system should be designed to allow regular updating and incorporation of security patches. |
| **Pentesting** | It is advisable to perform regular penetration tests of the system's software and hardware to identify possible vulnerabilities and implement corrective measures. |
| **Resilient design** | The design of the system should be as resilient as possible to adversities such as power outages or loss of connectivity. |
| **Privacy by design** | |
| **Privacy Impact Assessment (PIA)** | Designers should identify, examine, and assess the risks that may exist in relation to the personal data collected by the system. They should (1) identify the categories of data to be processed and (2) the type of processing they will be subject to, (3) identify, assess, and prioritise potential risks and threats, (4) adopt appropriate mitigation measures and (5) monitor the data flow. The use of the LINDDUN methodology based on data flow diagrams is strongly recommended. |
| **Data Minimisation** | The amount of personal data collected should be reduced as much as possible. Only data that is strictly necessary for the service and operation of the system should be collected. |
| **Access controls** | Restrictive permissions should be implemented so that only authorised personnel have access to files containing personal or sensitive data. |
| **Default login** | All system user accounts should require login by username and password or other verification and authentication mechanisms. |
| **Strong passwords** | Passwords should be secure and strong. The use of common or overly simple passwords should be avoided. They should be at least 8 characters long, including upper- and lower-case letters, numbers and special characters. The system should require that passwords have these characteristics and that they are changed with some frequency. Where possible, passwords should be audited to ensure they are secure. |
| **Double authentication factor** | Accounts with special privileges (e.g. access to large amounts of personal and/or sensitive data) should be protected by a double authentication factor. |
| **Anonymisation of personal data** | Personal Identifiable Information should be removed from the dataset as far as possible. This can be done by direct deletion, masking, generalisation or tokenisation. |
| **Encryption of data and communications** | Appropriate and robust cryptographic techniques should be used to ensure privacy and confidentiality of data and communications. |
| **Data retention and disposal policies** | Clear privacy policies should be implemented during design. They should define what data is collected, what data is retained and for how long. |

| | |
|---|---|
| **User awareness** | System users should be made aware of privacy risks and threats, as well as the corrective and self-protective measures they can take to mitigate them. |
| **Artificial Intelligence** | |
| **Decision-making** | Users should be the sole decision-makers, with the system merely a supporting tool. |
| **Security of AI systems** | AI systems should be protected against previously identified risks and threats. Containment plans should be in place in the event of adverse effects. |
| **Data accuracy** | The data on which the algorithm is trained must be accurate to avoid personal and/or material damage. Where possible, biases introduced into the system should be identified and eliminated. It is advisable to work with a culturally and academically diverse team to avoid bias. |
| **Privacy** | Protocols should be established about who can access what data and in what situations. Measures must be taken to anonymise, aggregate and encrypt data. It must be possible for humans to audit the data. |
| **Explainability and transparency of processes** | Users should know at all times that they are interacting with an AI system and, as far as possible, be able to understand the rationale behind the decisions it recommends. The objectives, limitations, benefits and risks of the system and its decisions should be communicated to the user. |
| **Environment and social welfare** | Identify, study, and evaluate the impact that the AI system may have on the environment (especially the issue of energy consumption), as well as on labour issues, institutions and democracy. |
| **Audits** | Systems should be designed in such a way that they can be audited. It should be possible to evaluate their data, their algorithms and their design process. |

# 6. LIST OF ABBREVIATIONS

| Abbreviation | Meaning |
|---|---|
| **ACRAI** | Austrian Council for Robotics and Artificial Intelligence |
| **AI** | Artificial Intelligence |
| **AIDA** | Artificial Intelligence and Data Act |
| **BIOS** | Basic Input/Output System |
| **CBDRM** | Community-based Disaster Risk Management |
| **CCTV** | Closed-circuit Television |
| **DFD** | Data Flow Diagram |
| **DL** | Deep Learning |
| **DRM** | Disaster Risk Management |
| **DSS** | Decision Support System |
| **EU** | European Union |
| **EU ETS** | European Union Emissions Trading System |
| **GDPR** | General Data Protection Regulation |
| **HLEG-AI** | High Level Group on Artificial Intelligence |
| **ICT** | Information and Communication Technologies |
| **IoT** | Internet of Things |
| **IPCC** | Intergovernmental Panel on Climate Change |
| **ML** | Machine Learning |
| **OECD** | Organization for Economic Cooperation and Development |
| **PIA** | Privacy Impact Assessment |
| **PII** | Personally Identifiable Information |
| **SCDT** | Smart City Digital Twin |

| | |
|---|---|
| **UAV** | Unmanned Aerial Vehicle |
| **UN** | United Nations |
| **UNESCO** | United Nations Educational, Scientific and Cultural Organization |

# 7. REFERENCES

Ahmad, K., Maabreh, M., Ghaly, M., Khan, K., Qadir, J., & Al-Fuqaha, A. (2022). Developing future human-centered smart cities: Critical analysis of smart city security, Data management, and Ethical challenges. *Computer Science Review*, 43, 100452. https://doi.org/10.1016/j.cosrev.2021.100452

Alamer, M., & Almaiah, M. A. (2021). Cybersecurity in Smart City: A systematic mapping study. *International Conference on Information Technology* (ICIT) pp. 719-724. DOI: 10.1109/ICIT52682.2021.9491123

Al-Turjman, F., Zahmatkesh, H., & Shahroze, R. (2022). An overview of security and privacy in smart cities' IoT communications. *Transactions on Emerging Telecommunications Technologies*, 33(3), e3677. https://doi.org/10.1002/ett.3677

Argyri, I., Kamara, A. M., & Pitsi, E. (2022) The new Law on "Emerging information and communication technologies, strenghtening digital governanve and other provisions" (Law 4961/2022) is here. *Kyriakides Georgopoulos Law Firm.* Available in: https://kglawfirm.gr/wp-content/uploads/2022/11/Emerging-Technologies-Law-4961-2022-c-6.pdf

Asif, M., Aziz, Z., Bin Ahmad, M., Khalid, A., Waris, H. A., & Gilani, A. (2022). Blockchain-Based Authentication and Trust Management Mechanism for Smart Cities. *Sensors*, 22(7), 2604. https://doi.org/10.3390/s22072604

Atlam, H. F. & Wills, G.B. (2020): "IoT Security, Privacy, Safety and Ethics." In Farsi, M., Daneshkhah, A., Hosseinian-Far, A. & Jahankhani, H. (eds.) *Digital Twin Technologies and Smart Cities.* Cham: Springer Nature Switzerland, pp. 123-149. https://doi.org/10.1007/978-3-030-18732-3

Bittner, S., Grabmaier, I., Kainz, S., Shtefchyk, K., Barrado, C., Zamora, D., Bakunc, S., Triantafyllou, I., Tsaloukidis, J. & Petropoulos, N. (2023) Deliverable 3.2. Report on Participatory Design Process. *PANTHEON Consortium.*

Bonney, M. S., de Angelis, M., Dal Borgo, M., Andrade, L., Beregi, S., Jamia, N. & Wagg, D. J. (2022) Development of a digital twin operational platform using Phyton Flask. *Data-Centric Engineering*, 3(e1). https://doi.org/10.1017/dce.2022.1

Brey, P. & Dainow, B. (2020) Ethics by Design and Ethics of Use approaches for Horizon Europe artificial intelligence projects. *SIENNA project.* Available in: https://sienna-project.eu/digitalAssets/915/c_915554-l_1-k_sienna-ethics-by-design-and-ethics-of-use.pdf

Burgess, P. & Kloza, D. (eds.) (2021) *Border Control and New Technologies*. Brussel: Academic and Scientific Publishers (ASP). DOI: 10.46944/9789461171375

Calvo, P. (2019) The ethics of Smart City (EoSC): moral implications of hiperconnectivity, algorithmization and the datafication of urban digital society. *Ethics and Information Technology.,* 22(1) pp. 141-149. https://doi.org/10.1007/s10676-019-09523-0

Chang, V. (2021) An Ethical Framework for Big Data and Smart Cities. *Technological Forecasting and Social Change*, vol. 165. https://doi.org/10.1016/j.techfore.2020.120559

*Charter of Fundamental Rights of the European Union* (2016a) *Official Journal of the European Union* C202, 7 June, pp.389-405. Available in: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:C:2016:202:FULL

Coccoli, J. (2017) The challenges of new technologies in the implementation of human rights: an analysis of some critical issues in the digital era. *Peace Human Rights Governance, 1*(2), 223-250. https://phrg.padovauniversitypress.it/2017/2/4

Coeckelbergh, M. (2020). *AI Ethics.* Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/12549.001.0001

*Consolidated versions of the Treaty on European Union* (2016b) *Official Journal of the European Union* C202, 7 June, pp.13-46. Available in https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:C:2016:202:FULL

*Consolidated version of the Treaty on the Functioning of the European Union* (2016c) *Official Journal of the European Union* C202, 7 June, pp.47-200. Available in: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:C:2016:202:FULL

Council of Europe (1950) Convention for the Protection of Human Rights and Fundamental Freedoms. In *Council of Europe Treaty Series* 005. Available in: https://rm.coe.int/1680063765

Cui, L., Xie, G., Qu, Y., Gao, L., & Yang, Y. (2018). Security and privacy in smart cities: Challenges and opportunities. *IEEE access*, 6, pp. 46134-46145. DOI: 10.1109/ACCESS.2018.2853985

Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council. Available in: https://eur-lex.europa.eu/eli/dir/2011/83/oj

Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources. Available in: https://eur-lex.europa.eu/eli/dir/2018/2001/oj

Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive). Available in: https://eur-lex.europa.eu/eli/dir/2022/2555/oj

Directive (EU) 2023/1791 of the European Parliament and of the Council of 13 September 2023 on energy efficiency and amending Regulation (EU) 2023/955. Available in: https://eur-lex.europa.eu/eli/dir/2023/1791/oj

European Commission (2020) White paper on Artificial Intelligence - A European approach to excellence and trust. Available in: https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

European Union Agency for Cybersecurity (ENISA) (2023) *Expression des Besoins et Identification des Objectifs de Sécurité (EBIOS).* Available in: https://www.enisa.europa.eu/topics/risk-management/current-risk/risk-management-inventory/rm-ra-methods/m_ebios.html

Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK) (2021) *Artificial Intelligence Strategy of the Austrian Federal Government. Artificial Intelligence Mission Austria 2023 (AIM AT 2023),* Vienna. Available in: https://digital-skills-jobs.europa.eu/sites/default/files/2023-10/Artificial%20Intelligence%20Mission%20Austria%202030_AIM_AT_2030.pdf

Helbing, D. & Argota Sanchez-Vaquerizo, J. (2023) Digital Twins: Potentials, Ethical Issues, and Limitations. *In:* Zwitter, A. & Gstrein, J. (Eds.): *Handbook on the Politics and Governance of Big Data and Artificial Intelligence.* Massachusetts: Edward Elgar Publishing, pp. 64-104. Available in: http://dx.doi.org/10.2139/ssrn.4167963

Helbing, D., Fanitabasi, F., Giannotti, F., Hänggli, R., Hausladen, C. I., van den Hoven, J., Mahajan, S., Pedreschi, D., & Pournaras, E. (2021) Ethics of Smart Cities: Towards Value-Sensitive Design and Co-Evolving City Life, *Sustainability,* 13(20). https://doi.org/10.3390/su132011162

High-Level Expert Group on Artificial Intelligence (HLEG-AI) (2019) *Ethics guidelines for Trustworthy AI.* Brussels: European Commission. Available in: https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

Hiller, J. S. & Blanke, J. M. (2017) Smart Cities, Big Data, and the Resilience of Privacy. *Hastings Law Journal*, 68(2), pp. 309-356. Available in: https://repository.uclawsf.edu/hastings_law_journal/vol68/iss2/3

Hiranandani, V. (2011) Privacy and security in the digital age: contemporary challenges and future directions. *The International Journal of Human Rights, 15*(7), pp. 1091-1106. DOI: https://doi.org/10.1080/13642987.2010.493360

Ingeno, J. (2018) "*Software Architect's Handbook: Become a successful software architect by implementing effective architecture concepts,"* Birmingham: Packt Publishing.

Intergovernmental Panel on Climate Change (IPCC) (2023) Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, 184 pp., doi: 10.59327/IPCC/AR6-9789291691647.

Kitchin, R., & Dodge, M. (2019). The (in) security of smart cities: Vulnerabilities, risks, mitigation, and prevention. *Journal of urban technology*, 26(2), pp. 47-65. https://doi.org/10.1080/10630732.2017.1408002

Kitchin, R. (2016) The ethics of smart cities and urban science. The ethics of smart cities and urban science. *Phil. Trans. R. Soc. A*., 374(20160115). https://doi.org/10.1098/rsta.2016.0115

Kouroutakis, A. E. (2019) The Architecture of the Emergency Framework of Greece: Inactivity and Second Generation Emergencies. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3333118

Kyoto Protocol to the United Nations Framework Convention on Climate Change, Dec. 10, 1997, 2303 U.N.T.S. 162. Available in: https://treaties.un.org/doc/Publication/UNTS/Volume%202303/v2303.pdf

Mahmoud, R., Yousuf, T., Aloul, F. & Zualkernan, I. (2015) Internet of Things (IoT) Security: Current Status, Challenges and Prospective Measures. In *2015 10th International*

*Conference for Internet Technology and Secured Transactions (ICITST).* DOI: 10.1109/ICITST.2015.7412116

Martín, Y. S. & Kung, A. (2018) Methods and Tools for GDPR Compliance through Privacy and Data Protection Engineering. *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pp. 108-11. DOI: 10.1109/EuroSPW.2018.00021

Meek, T., Barham, H., Beltaif, N., Kadoor, A., & Akhter T. (2016) Managing the ethical and risk implications of rapid advances in Artificial Intelligence: a literature review. *2016 Portland International Conference on Management of Engineering and Technology (PICMET).* DOI: 10.1109/PICMET.2016.7806752

Mitsios, S. (2022) L. 4961/2022: The Greek Legal Framework on Emerging Technologies. *Ernst & Young.* Available in: https://www.ey.com/en_gr/tax/tax-alerts/l-4961-2022-the-greek-legal-framework-on-emerging-technologies

Mohamed, N., Al-Jaroodi, J., Jawhar, I., Idries, A. & Mohammed, F. (2018) Unmanned aerial vehicles applications in future smart cities. *Technological Forecasting & Social Change.* https://doi.org/10.1016/j.techfore.2018.05.004

National Institute of Standards and Technology (NIST) (2023) *Secure Software Development Framework.* Available in: https://csrc.nist.gov/projects/ssdf

Organization for Economic Cooperation and Development (OECD) (2023) *G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 common understanding on generative AI. Report prepared for the 2023 Japanese G7 presidency and the G7 digital and tech working group.* OECD Publishing. Available in: https://www.oecd.org/publications/g7-hiroshima-process-on-generative-artificial-intelligence-ai-bf3c0c60-en.htm

Organization for Economic Cooperation and Development (OECD) (2019) *Recommendation of the Council on Artificial Intelligence,* OECD/LEGAL/0449. Available in: https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449

Paris Agreement to the United Nations Framework Convention on Climate Change, Dec. 12, 2015, T.I.A.S. No. 16-1104. Available in: https://unfccc.int/sites/default/files/english_paris_agreement.pdf

Pandey, P., Golden, D., Peasley, S., & Kelkar, M. (2019) Making smart cities cybersecure: ways to address distinct risks in an increasingly connected urban future. *Deloitte Development LL.* Available in: https://www2.deloitte.com/us/en/insights/focus/smart-city/making-smart-cities-cyber-secure.html

Perry, S. & Roda, C. (2017) *Human rights and digital technology. Digital tightrope.* London: Palgrave Macmillan. https://doi.org/10.1057/978-1-137-58805-0

Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. COM/2021/206 final. Available in: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

Regulation (EC) No 765/2008 of the European Parliament and of the Council of 9 July 2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products and repealing Regulation (EEC) No 339/93. Available in: https://eur-lex.europa.eu/eli/reg/2008/765/oj

Regulation (EU) 2017/2394 of the European Parliament and of the Council of 12 December 2017 on cooperation between national authorities responsible for the enforcement of consumer protection laws and repealing Regulation (EC) No 2006/2004. Available in: https://eur-lex.europa.eu/eli/reg/2017/2394/oj

Regulation (EU) 2016/679 of the European Parliament and of the Council, of April 27, 2016, relating to the protection of natural persons with regard to the processing of personal data and the free circulation of these data and by which repeals Directive 95/46/EC (General Data Protection Regulation) Available in: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex%3A32016R0679

Righetti, F., Vallati, C., & Anastasi, G. (2018) IoT Appliacionts in Smart Cities: a Perspective into Social and Ethical Issues. In *2018 IEEE International Conference on Smart Computing (SMARTCOMP),* pp. 387-392, doi: 10.1109/SMARTCOMP.2018.00034.

Robles-González, A., Parra-Arnau, J. & Forné, J. (2020) A LINDDUN-Based framework for privacy threat analysis on identification and authentication processes. *Computers & Security*, 94(101755). https://doi.org/10.1016/j.cose.2020.101755

Schoenfield, B. S. E. (2015) *Securing Systems: applied security architecture and threat models.* New York: CRC Press.

Sharma, S. (2020) *Data privacy and GDPR handbook.* New Jersey: John Wiley & Sons. DOI:10.1002/9781119594307

Shtefchyk, K. & Vallès, L. (2023) *Deliverable 10.3. AI – Requirement No. 3.* PANTHEON Consortium.

Stern, D.I. & Kaufmann, R.K. (2014) Anthropogenic and natural causes of climate change. *Climatic Change,* vol. 122, pp. 257-269. https://doi.org/10.1007/s10584-013-1007-x

Stöger, K. (2021) 'Austria: Legal Response to Covid-19', in Jeff King & Octávio LM Ferraz et al. (eds), *The Oxford Compendium of National Legal Responses to Covid-19.* DOI: 10.1093/law-occ19/e28.013.28

Syifa, M., Kadavi, P. R. & Lee, C.W. (2019) An Artificial Intelligence application for post-earthquake damage mapping in Palu, Central Sulaweis, Indonesia. *Sensors,* 19(3). https://doi.org/10.3390/s19030542

The Constitution of Greece [Greece], 18 April 2001. Available in: https://www.refworld.org/docid/4c52794f2.html

The Federal Constitutional Law of 1920 as amended in 1929 as to Law No. 153/2004, December 30, 2004. Available in: https://constitutionnet.org/sites/default/files/Austria%20_FULL_%20Constitution.pdf

Toh, C. K. (2020) Security for smart cities. *IET Smart Cities,* 2(2), pp. 95-104. https://doi.org/10.1049/iet-smc.2020.0001

Tsaloukidis, J., Petsioti, P., Gatsogianni, C., Georgiou, E., Stergiopoulos, G., Karamousadakis, M., Apostolopoulou, V, Nakos, S., Bularca, O., Barrado, C. & Sharples, J. (2023). Deliverable

3.1. PANTHEON Technology Roadmap for Disaster Resilient Communities. *PANTHEON Consortium.*

Tzafestas, S. G. (2018) Ethics and Law in the Internet of Things World. *Smart Cities,* 1(1), pp. 98-120. https://doi.org/10.3390/smartcities1010006

United Nations (1948) Universal Declaration of Human Rights. Available in: https://www.un.org/en/about-us/universal-declaration-of-human-rights

United Nations (General Assembly). (1966a). International Covenant on Civil and Political Rights. *Treaty Series*, 999, 171. Available in: https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-4&chapter=4&clang=_en

United Nations (General Assembly). (1966b). International Covenant on Economic, Social, and Cultural Rights. *Treaty Series*, 999, 171. Available in: https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-3&chapter=4&clang=_en

United Nations (General Assembly). (1968). Human rights and scientific and technological developments, *1748th plenary meeting*. Available in: https://digitallibrary.un.org/record/202686?ln=en

United Nations (General Assembly). (1979) Convention on the Elimination of All Forms of Discrimination Against Women, *Treaty Series*, vol. 1249 Available in: https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-8&chapter=4&clang=_en

United Nations (General Assembly). (1989) Convention on the Rights of the Child, *Treaty Series*, vol. 1577. Available in: https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child

United Nations (General Assembly). (2006). Convention on the Rights of Persons with Disabilities. *Treaty Series*, 2515. Available in: https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-persons-disabilities

United Nations (Human Rights Council) (2014) The right to privacy in the digital age: Report of the Office of the United Nations High Commissioner for Human Rights. A/HRC/27/37. Available in: https://documents-dds-ny.un.org/doc/UNDOC/GEN/G14/088/54/PDF/G1408854.pdf?OpenElement

United Nations (2020). *Report of the Secretary-General. Roadmap for Digital Cooperation*. Available in: https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf

United Nations (Chief Executive Board for Coordination) (2022) *Principles for the ethical use of artificial intelligence in the United Nations system.* CEB/2022/2/Add.1. Available in: https://unsceb.org/session-report-389

United Nations (Secretary-General) (2023a) *Secretary-General's statement at the UK AI Safety Summit.* Available in: https://www.un.org/sg/en/content/sg/statement/2023-11-02/secretary-generals-statement-the-uk-ai-safety-summit#:~:text=The%20speed%20and%20reach%20of,governance%20is%20wide%20and%20growing.

United Nations (Secretary-General) (2023b) *Press remarks launching high-level advisory body on Artificial Intelligence.* Available in: https://www.un.org/sg/en/content/sg/speeches/2023-10-26/secretary-general%E2%80%99s-remarks-press-conference-launching-high-level-advisory-body-artificial-intelligence%C2%A0

United Nations Educational, Scientific and Cultural Organizations (UNESCO) (2022) *Recommendation on the Ethics of Artificial Intelligence.* SHS/BIO/PI/2021/1. Paris: UNESCO. Available in: https://unesdoc.unesco.org/ark:/48223/pf0000381137

Wheeler, D. M., Wheeler, J.C. & Fagbemi, D. D. (2020) *The IoT architect's guide to attainable security and privacy*. New York: CRC Press. DOI: 10.1201/9780367440930

Yigitcanlar, T., Desouza, K., Butler, L. & Roozkhost, F. (2020) Contributions and risks of Artificial Intelligence (AI) in building smarter cities: insights from a systematic review of the literature. *Energies,* 13(6). https://doi.org/10.3390/en13061473

Ziosi, M., Hewitt, B., Juneja, P., Taddeo, M. & Floridi, L. (2022) Smart cities: reviewing the debate about their ethical implications, *AI & Society*. https://doi.org/10.1007/s00146-022-01558-0